4

PA 228560

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

March 30, 2000

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: *60/119,388*
FILING DATE: *February 10, 1999*

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

H. PHILLIPS
Certifying Officer

37 PRO

FOLEY & LARDNER
3000 K Street, N.W., Suite 500
P.O. Box 25696
Washington, D.C. 20007-8696
(202) 672-5300

PROVISIONAL APPLICATION FOR PATENT

Assistant Commissioner for Patents
Washington, D. C. 20231

Sir:

This is a request for filing a **PROVISIONAL APPLICATION FOR PATENT** UNDER 37 CFR 1.53(c).

INVENTOR(S)/APPLICANT(S)			
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (City & either State or Country)
MEMELINK	Johan		Leiden, Holland
VAN DER FITS	Leslie		Hoogmade, Holland
MENKE	Frank	L.H.	Leiden, Holland
KIJNE	Jan	W.	Leiden, Holland

TITLE OF THE INVENTION
METHOD OF MODULATING METABOLITE BIOSYNTHESIS IN RECOMBINANT CELLS

In connection with this application, the following are enclosed:

65 Pages of Specification (Optional: ☒ Abstract ☒ Claims 128) including 5 pages of Sequence Listing.

12 Sheets of Drawings

___ Assignment to: ___

___ Statement of Small Entity Status

XX Other: Check for $150.00.

Attorney Docket No. 030307/0167

The fee has been calculated as shown below. (Small entity fees indicated in parentheses.)

Filing Fee	$150 ($75)	$150.00
Rule 17(k) fee for non-English text	$130	
Assignment Recording Fee	$ 40	
	TOTAL FEE	$150.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No ☐ Yes, the name of the U.S. Government agency and the Government contract number are: .

A check in the amount of the above TOTAL FEE is attached. The Commissioner is hereby authorized to charge any deficiency or credit any overpayment to Deposit Account No. 19-0741.

Respectfully submitted,

Date: February 10, 1999
Docket No.: 030307/0167

Stephen A. Bent
Reg. No. 29,768

METHOD OF MODULATING METABOLITE BIOSYNTHESIS IN RECOMBINANT CELLS

FIELD OF THE INVENTION

The present invention relates to the field of modulating the biosynthesis of metabolites in cells. More particularly, the invention relates to the cloning and expression of transcription factor genes, such as genes coding for AP2 domain class transcription factors and the use of such transcription factor for regulating the expression of genes involved in the biosynthesis of metabolites or precursors therefor. Especially the invention relates to the cloning and expression of AP2 domain class plant transcription factors and use hereof for regulating the expression of biosynthetic genes involved in the secondary metabolite biosynthesis in plant cells.

BACKGROUND OF THE INVENTION AND PRIOR ART

Generally, two basic types of metabolites are synthesised in cells, i.e. those referred to as primary metabolites and those referred to as secondary metabolites. A primary metabolite is any intermediate in, or product of the primary metabolism in cells. The primary metabolism in cells is the sum of metabolic activities that are common to most, if not all, living cells and are necessary for basal growth and maintenance of the cells. Primary metabolism thus includes pathways for generally modifying and synthesising certain carbohydrates, proteins, fats and nucleic acids, with the compounds involved in the pathways being designated primary metabolites.

In contrast hereto, secondary metabolites usually do not possess a basal function in cell growth and maintenance. They are a group of chemically very diverse products that often have a restricted taxonomic distribution. Secondary metabolites normally exist as members of closely related chemical families, usually of a molecular weight of less than 1,500, although some bacterial toxins are considerably larger. Two examples of fungal cell secondary metabolites are penicillin and ergotamine.

Plant metabolites include a diverse array of chemically unrelated compounds such as carbohydrates and lipids (e.g. mono-, oligo- and polysaccharides, sugar alcohols, organic

acids, fatty acids and lipids, acetylenes and thiophenes), nitrogen-containing compounds (e.g. amino acids, amines, glycosides, glucosinolates, purines, pyrimidines and polypeptides) of which most, but not all, generally are referred to as primary metabolites. Accordingly, some compounds such as fatty acids, sugars and steroids may e.g. be categorised both as primary metabolites and secondary metabolites (see e.g. Dewick, P.M., 1997, Medicinal Natural Products A Biosynthetic Approach, John Wiley & Sons, Chichester).

Secondary plant metabolites include e.g. alkaloid compounds (e.g. terpenoid indole alkaloids and indole alkaloids), phenolic compounds (e.g., quinones, lignans and flavonoids), terpenoid compounds (e.g. monoterpenoids, iridoids, sesquiterpenoids, diterpenoids and triterpenoids). In addition, secondary metabolites include small molecules (i.e., having a molecular weight of less than 600), such as substituted heterocyclic compounds which may be monocyclic or polycyclic, fused or bridged.

Many plant secondary metabolites have value as pharmaceuticals, food colours, flavours and fragrances. Plant pharmaceuticals include e.g.: taxol, digoxin, colchicine, codeine, morphine, quinine, shikonin, ajmalicine and vinblastine. Examples of secondary metabolites that are useful as food additives include anthocyanins, vanillin, and a wide variety of other fruit and vegetable flavours and texture modifying agents.

Plant secondary metabolites such as terpenoid indole alkaloids (TIA) represent a class of pharmaceutically useful compounds which naturally occur in many plant species.

Some plant secondary metabolites are linked to plant or plant cell defence mechanism and may e.g. confer to the plant antimicrobial activity, protection against UV light, herbivores, pathogens, insects and nematodes, and the ability to grow at low light intensity.

There are numerous examples of the application of plant secondary metabolites such as TIA's in medicine. The monomeric alkaloids serpentine and ajmalicine found in *Catharanthus roseus* are e.g. used as tranquillisers and to reduce hypertension, respectively. The dimeric alkaloids vincristine and vinblastine, also found in *C. roseus*, are potent antitumor drugs. Camptothecin, a monomeric TIA found in *Camptotheca acuminata*, also possesses anti-tumor activity. Quinine from *Cinchona officinalis* is used in malaria treatment.

However, a major problem associated with the industrial use of the above metabolites is that only very small or variable amounts of these compounds are present in plants. The recovery of useful metabolites from their natural sources is thus in many instances difficult due to the enormous amounts of source material which may be required for the isolation of utilisable quantities of the desired products. As an example, over 500 kg of *Catharanthus roseus* is needed to obtain 1g of vincristine. Extraction is both costly and tedious, requiring large quantities of raw material and extensive use of chromatographic fractionation procedures. The low levels of these compounds in plants may also imply that many of the compounds are not detected when performing normal screening procedures, hence many unknown compounds may exist.

Biosynthesis of TIA compounds proceeds in most plants via many enzymatic steps. TIA compounds consist of an indole moiety provided by tryptamine and a terpenoid portion provided by the iridoid glucoside compound secologanin. Tryptamine is derived from primary metabolism by a single enzymatic conversion of the amino acid tryptophan, a reaction catalysed by the enzyme tryptophan decarboxylase (TDC) (EC 4.1.1.28). The biosynthesis of secologanin requires a number of enzymatic conversions of which the first step is the hydroxylation of geraniol by the enzyme geraniol 10-hydroxylase (G10H). Tryptamine and secologanin are condensed by the enzyme strictosidine synthase (STR) (EC 4.3.3.2) to form strictosidine, which is the general precursor of all terpenoid indole alkaloids found in plants. The first enzymatic conversion of strictosidine into strictosidine aglucon is catalysed by strictosidine-ß-D-glucosidase (SG) (EC 3.2.1.105). Many different enzyme activities convert strictosidine aglucon into the large variety of terpenoid indole alkaloid end products.

The best progress on molecular characterisation of the terpenoid indole alkaloid pathway has been made with *C. roseus* or Madagascar periwinkle, a member of the *Apocynaceae* family. *C. roseus* cells have the genetic potential to synthesise over a hundred terpenoid indole alkaloids. The biosynthesis of terpenoid indole alkaloids is strongly regulated, and depends on plant cell type and environmental conditions. Their biosynthesis is e.g. induced by fungal elicitors (Moreno et al., 1995), jasmonates (Gantet et al., 1998) and auxin starvation (Gantet et al., 1998). Many monomeric TIA compounds are found in all plant organs, but vindoline and vindoline-derived dimeric alkaloids are only found in chloroplast-containing plant tissues (Meijer et al., 1993).

Until now, only a limited number of genes coding for enzymes involved in terpenoid indole alkaloid biosynthesis have been cloned, such as cDNA clones encoding STR and TDC isolated from *C. roseus*. In addition, cDNA clones for NADPH:cytochrome P450 reductase (CPR) (Meijer et al., 1993), which is essential for the G10H-catalysed reaction, and SG have been isolated (Geerlings, Memelink, van der Heijden, and Verpoorte, unpublished results).

Gene expression studies by the present inventors have shown that the regulation of terpenoid indole alkaloid biosynthesis is controlled largely, if not uniquely, at the level of the expression of biosynthetic genes. Accordingly, analysis of the expression of the terpenoid indole alkaloid biosynthetic genes *Tdc* and *Str*1 showed that their expression is low, especially in cell cultures. It was found, that the level of gene expression is likely to be limiting for alkaloid production. Overexpression of a single biosynthetic gene (*Tdc* or *Str*1) in transgenic *C. roseus* cells resulted in significantly elevated levels of the corresponding enzyme activity. However, this did not result in elevated alkaloid levels, presumably because many other enzymes are limiting and would need to be overexpressed.

These studies have further demonstrated what is already known, that the genes are coordinately regulated depending on cell type or environmental conditions. *Str*1 and *Tdc* mRNA accumulate in suspension-cultured cells after auxin starvation (Pasquali et al., 1992) or phosphate starvation, exposure to fungal elicitors (Pasquali et al., 1992) or (methyl) jasmonate (Menke et al., 1999) and their distribution in the plant is developmentally regulated with the highest levels in the roots (Pasquali et al., 1992). In leaves *Tdc* and *Str*1 are induced by a UV-B light pulse (Ouwerkerk and Memelink, unpublished results). *Cpr* mRNA accumulation is rapidly induced by fungal elicitor (Meijer et al., 1993).

The observations that mRNA from genes involved in the biosynthesis of TIA compounds such as *Tdc* and *Str*1 mRNAs coordinately accumulate in response to fungal elicitors, (methyl)-jasmonate, UV light, auxin depletion, phosphate depletion, and have similar tissue-specific distributions, have led the present inventors to hypothesise that the *Tdc* and *Str*1 genes might be controlled by one or more common regulating factor(s) or substance(s). It was further hypothesised that among possible regulating factors, transcription factors could have such a regulating effect.

It is known that certain transcription factors can regulate complex cell differentiation processes in animals involving numerous target genes. A notable example is muscle differentiation, where either one of a set of myogenic bHLH transcription factors (MyoD, myogenin, Myf5, MRF4) in combination with the MADS-domain transcription factor MEF2 induces muscle cell differentiation and switches on numerous muscle-specific genes. Other examples include homeodomain transcription factors in the fruit fly that regulate cell processes resulting in the determination of segment identity.

The pathway that is most extensively studied in plant secondary metabolism at the transcriptional level, is the one leading to the formation of the anthocyanin pigments. The genes encoding flavonoid biosynthetic enzymes are controlled by a combination of two distinct transcription factor species, one of which has homology to the protein encoded by the vertebrate proto-oncogene c-*Myb*, and the other with the vertebrate bHLH protein encoded by the proto-oncogene c-*Myc*. These transcription factors bind to specific sequences in the promoters of the target genes (Martin and Paz-Ares, 1997).

The DNA-binding domain of plant MYB proteins consists of two, or for some of them, one imperfect repeat(s) (Martin and Paz-Ares, 1997). The MYC proteins have a bHLH-type (basic helix-loop-helix) DNA-binding domain and recognise variants of the sequence CANNTG. About ten enzymes are involved in the biosynthesis of anthocyanins starting from phenylalanine. In maize, the entire set of genes encoding these enzymes are thought to be regulated coordinately by the *Myc* gene *R* and the *Myb* gene *C1* in the aleurone (epidermal layer of the kernel endosperm), and by homologous genes in other parts of the plant (Holton and Cornish, 1995).

Overexpression of the maize *Lc* gene encoding a MYC-type regulatory protein in *Petunia* upregulated the whole flavonoid biosynthetic pathway starting from *Chs* and including the earlier and later genes, resulting among others in intensely pigmented leaves (Bradley et al., 1998). The expression of the general phenylpropanoid genes *Pal* and *C4h* was not affected by *Lc* overexpression, indicating that *Lc* only regulates structural genes in the flavonoid branch.

Introduction of the maize *R* and *C1* MYC-type regulators in *Arabidopsis* intensified pigmentation in normally pigmented tissues and induced pigmentation in plant tissues that

are normally unpigmented (Lloyd et al., 1992). In maize cell suspension, ectopic expression of *C1* and *R* led to the accumulation of anthocyanins and a number of other related 3-hydroxy flavonoids (Grotewold et al., 1998). In addition, six anthocyanin structural genes that are targets for *C1/R* were expressed at high levels in the transgenic cell line.

In the plant *Arabidopsis thaliana*, it has been shown that overexpression of the transcription factor CBF-1 that belongs to the AP2 domain class transcription factors resulted in coordinate upregulation of a set of cold-regulated genes (Jaglo-Ottosen et al., 1998). However, there are no suggestions that transcription factors of this class may have an effect on the biosynthesis of metabolites in plant cells and other cells.

The present inventors have now discovered that transcription factors having an AP2 DNA-binding domain are highly useful as central regulators of complex metabolite pathways involving numerous target genes for such transcription factors. This discovery has opened up for providing novel effective means of generating novel metabolite compounds, significantly enhancing the yield of commercially valuable metabolite compounds and also for enhancing the tolerance of plants towards exogenous stress factors and conditions.

SUMMARY OF THE INVENTION

Accordingly, the invention relates in one aspect to a method of modulating in a cell the expression of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor, the method comprising inserting into the cell a nucleotide sequence coding for a transcription factor other than a transcription factor having a bHLH-type or a MYB-type DNA-binding domain, the nucleotide sequence is operably linked to at least one expression regulating sequence, and/or modifying the expression of a nucleotide sequence coding for such a transcription factor already present in the cell, the transcription factor is capable of regulating the expression of at least one gene coding for a gene product involved in the biosynthesis of said metabolite or precursor, and subjecting the cell to conditions where the inserted nucleotide sequence coding for a transcription factor is expressed or the expression of the nucleotide sequence already present in the cell is modulated.

In a further aspect the invention relates to a method of modulating the stress resistance of a cell, the method comprising inserting into the cell a nucleotide sequence coding for a transcription factor, the nucleotide sequence is operably linked to at least one expression regulating sequence, and/or modifying the expression of a nucleotide sequence coding for such a transcription factor already present in the cell, the transcription factor is capable of regulating the expression of at least one gene coding for a gene product involved in the biosynthesis of said metabolite or precursor, and subjecting the cell to conditions where the inserted nucleotide sequence coding for a transcription factor is expressed or the expression of the nucleotide sequence already present in the cell is modulated, the expression of said nucleotide sequence resulting in a modified responsiveness of the cell towards exogenous stress conditions.

There is also provided a recombinant cell having, relative to its parent cell, enhanced or reduced biosynthesis of a metabolite or a precursor therefor, and/or enhanced or reduced expression of a gene product involved in metabolite production, the cell comprising a nucleotide sequence coding for a transcription factor other than a transcription factor having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating the expression of at least one gene coding for a gene product involved in the metabolite production, said sequence is inserted into the cell and/or its expression is modified by operably linking it to a regulating sequence with which it is not natively associated.

The invention pertains in a still further aspect to a method of producing a metabolite including a plant secondary metabolite, the method comprising providing a recombinant cell according to the invention, cultivating the cell under conditions where the nucleotide sequence coding for the transcription factor regulating the expression of at least one gene coding for a gene product involved in the metabolite production is expressed, and recovering the metabolite.

There is also provided a method of constructing a recombinant cell according to the invention, the method comprising the steps of (i) identifying in a source cell a nucleotide sequence coding for a transcription factor other than a transcription factor having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating in the source cell the expression of at least one gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, (ii) isolating said nucleotide sequence

and (iii) inserting said isolated nucleotide sequence into a host cell comprising a gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, the insertion of said nucleotide sequence resulting in that the expression of the gene of the host cell is modulated relative to the expression level in a host cell not comprising the inserted nucleotide sequence.

In still further aspects there are provided an isolated nucleic acid molecule comprising a nucleotide sequence coding for an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7, a polypeptide which includes an amino acid sequence having at least 40% homology including at least 50%, such as at least 60%, at least 70%, at least 80% and at least 90% to an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7 and a polypeptide having the amino acid sequence shown in SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7.

In other aspects, the invention pertains to a vector comprising a nucleic acid molecule according to the invention, a host cell comprising such a vector, and a method of isolating a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating in a cell the expression of at least one gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, the method comprising the steps of (i) probing a cDNA or genomic library with a probe comprising at least a fragment of nucleic acid according to claim 117 or 118; (ii) identifying a DNA clone that hybridises with said nucleic acid; and (iii) isolating the DNA clone identified in step (ii) wherein the nucleic acid sequence is coding for all or a part of said transcription factor, a method of isolating a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating in a cell the expression of at least one gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, the method comprising the steps of (i) synthesising an oligonucleotide primer set corresponding to at least a fragment of a sequence selected from the group consisting of SEQ ID NO:1, 2, 3 and 4; and (ii) amplifying cDNA or genomic DNA using said primer set in a polymerase chain reaction wherein the amplified nucleic acid fragment is coding for all or a part of said transcription factor to obtain a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain.

The invention also pertains to an isolated nucleic acid which is a nucleotide sequence as shown in SEQ ID NO: 8 and a mutant, allele or variant hereof and the use of this isolated nucleic acid sequence for isolating a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating in a cell the expression of at least one gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor.

DETAILED DISCLOSURE OF THE INVENTION

A major objective of the present invention is to provide a method of modulating in a cell the expression of one or more genes involved in the biosynthesis of a metabolite or a precursor for the metabolite. As mentioned above, the method comprises that a nucleotide sequence coding for a transcription factor other than a transcription factor having a bHLH-type or a MYB-type DNA-binding domain is inserted into the cell such that the nucleotide sequence is operably linked to at least one expression regulating sequence, and/or that the expression of a nucleotide sequence coding for such a transcription factor already present in the cell is modified.

In the present context, the term "modulating" refers to any change in the level of expression of a gene involved in the biosynthesis of a metabolite or a precursor for the metabolite relative to the level of expression in a corresponding cell type not having a modulated expression of the same gene or genes, i.e. in the wild type state. It will be appreciated that "modulating" may imply both reduced and enhanced levels of expression relative to that found in the wild type cell.

As used herein the term "cell" encompasses any eukaryotic or prokaryotic cell. Thus, cells include bacterial cells, fungal cells including yeast cells, animal cells, dictyostelium cells, algae cells, and plant cells. The cell may be in the form of a single, isolated cell or it may in the form of multicellular tissues including an animal or a plant organ, plant callus tissues and an entire animal or plant.

The expression "nucleotide sequence coding for a transcription factor" is used herein encompasses any nucleotide sequence coding for a transcription factor which does not comprise a bHLH-type or a MYB-type DNA-binding domain. It will be appreciated that such

a coding sequence may consist solely of coding DNA (ORFs) or, when it is derived from a eukaryotic cell, it may include introns.

In this context, the expression "transcription factor" refers to a protein that in a sequence-specific manner binds to DNA by recognising specific elements located in the promoter and/or enhancer regions of the corresponding genes. The binding of such transcription factors to the specific DNA elements modulates the activity of other components of the transcription machinery including basal transcription factors and RNA polymerase, and thereby positively or negatively modulates the initiation of mRNA synthesis. Transcription factors have in common that they comprise DNA binding domains capable of binding to specific elements of their target genes.

Transcription factors are classified in families based on conserved features of their DNA-binding domain. Within each family, DNA-binding domains are similar, but not identical. Small differences in amino acid sequences can cause significant differences in the DNA sequence requirement for binding. As a result of these differences within a family, it is impossible to reliably predict whether a certain DNA sequence will bind a certain transcription factor family member with high affinity, and conversely, based on the amino acid sequence of a certain transcription factor family member it is impossible to reliably predict which sequence it will bind with high affinity.

Transcription factors have a modular structure, consisting of domains with specialised functions, such as sequence-specific DNA binding, transcriptional activation, interaction with themselves or other proteins. These domains are in many cases functional even when fused to heterologous proteins. ORCA proteins contain the AP2-type DNA-binding domain. Since they activate transcription, it can be concluded that they contain an activation domain. It is conceivable that in e.g. *Catharanthus* cells the activity of AP2 domain class transcription factors is modulated via post-translational modifications (phosphorylation, myristoylation etc.) via interaction with other proteins, or combinations thereof. The activity of AP2 domain class transcription factors can be modulated using recombinant DNA techniques. The objectives could be to make the activity of the transcription factors independent from post-translational modifications, to enhance their activity, or to make the activity dependent on convenient inducing conditions. These objectives can be accomplished via single amino acid modifications, deletions of several amino acids, fusions with other proteins or protein domains, or combinations thereof. AP2 domain-containing

transcription factors may be constitutively targeted to the nucleus by addition of a nuclear localisation signal. Strong activation domains could be fused to the transcription factors in addition to or as a replacement for existing activation domains. Such activation domains include the Herpes simplex virus VP16 domain, the GAL4 activation domain, or synthetic activation domains known to be active in plants. Inducibility of transcription factor activity can be achieved via fusion with mammalian steroid hormone receptor domains. The resulting chimera can be activated by steroid hormone addition.

In accordance with the above method of the invention, the nucleotide sequence coding for the transcription factor is operably linked to at least one expression regulating sequence. In the present context, the term "operably linked" refers in general to a situation where a first nucleic acid sequence is linked with a second nucleic acid sequence and the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence (i.e. that the coding sequence is under the transcriptional control of the promotor). Generally, operably linked nucleic acid sequences are contiguous and, where necessary to join two protein coding regions, in reading frame. However, operably linked nucleic acid sequences may also be in a functional relationship without being contiguous. Typical examples of "expression regulating sequences" include promoter sequences, enhancer sequences, translation leader sequences, sequences coding for anti sense RNA and 3' non-coding sequences.

It will be understood that the term "promoter" includes both constitutive promoters and regulatable promoters. In respect of the latter type of promoters, the regulation is typically conferred by exogenous (extrinsic) or endogenous (intrinsic) factors including physical factors such as the temperature or light, chemical factors including the pH, the presence of inducer substances, absence/presence of nutrient compounds in the cell environment.

In one presently preferred embodiment, the cell in which the expression of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is a plant cell.

The term "plant cell" as used herein include protoplasts, fused cells, gamete producing cells, undifferentiated and differentiated cells, and cells which regenerate into whole plants. Accordingly, a seed comprising multiple plant cells capable of regenerating into a whole plant, is included in the definition of "plant cell". In useful embodiments, the plant cell is in

the form of a cell culture. As used herein, the term "plant" refers to any organism of the kingdom Plantae, including Bryophytes, Pteridophytes, Spermatophytes (*Gymnosperms* and *Angiosperms*, including monocotyledonous and dicotyledonous), thallophytes (including algae such as *Chrysophyta*, *Pyrrophyta*, *Euglenophyta*, *Rhodophyta*, *Phaeophyta* and *Clorophyta*), a whole plant, any plant part or propagule thereof, seed or hybrid progeny and descendants, a plant cell, a multiplicity of plant cells such as e.g. plant tissues, and plant cells in the form of a cell culture. Plantlets are also included within the meaning of "plant".

In specific embodiments, the method relates to plant cells of a species of the *Gentianales* order or the *Cornales* order such as a species selected from an indole alkaloid producing species. Such species include species of the *Apocynaceae*, *Alangiaceae, Loganiaceae, Icacinaceae, Cornaceae* and *Rubiaceae* families. One particularly interesting genus of the *Apocynaceae* family is the genus *Catharanthus* including the species *C. roseus, C. coriaceus, C. lanceus, C. longifolius, C. ovalis, C. pusillus, C. scitulus and C. trichophyllus*.

In accordance with the invention, the nucleotide sequence coding for a transcription factor as defined above can be derived from any cell type comprising such a sequence. Particularly useful sources for such sequences are plant cells of any of the above types and forms.

In one interesting embodiment, the nucleotide sequence coding for a transcription factor is a sequence coding for a transcription factor comprising at least one AP2 domain. Transcription factors comprising such a DNA-binding domain are generally referred to as AP2/EREBP transcription factors which constitute a class of transcription factors that at present have only been found in plants. They posses two features that are typical for transcription factors, namely sequence-specific DNA binding capability and ability to activate (or repress) transcription. (Riechmann and Meyerowitz, 1998).

Sequence similarity among the AP2/EREBP proteins is mostly limited to the AP2 DNA-binding domain (APETALA2 domain) itself (approximately 68 aa), which is the only region conserved among all proteins of the family. The distinguishing characteristic of proteins of the AP2/EREBP family is that they contain either one or two AP2 domains. Thus this family can be divided into two subfamilies based on whether the proteins contain one or two AP2 domains, the EREBP subfamily and the AP2 subfamily, respectively. The AP2 domains in

the AP2 subfamily are more related among members of this subfamily than to the AP2 domains of proteins of the EREBP subfamily and *vice versa*. However, many characteristics are common to the AP2 domains from both subfamilies. Two conserved segments are found within each AP2 domain, which have been referred to as the YRG element, and the RAYD element. The amino terminal part of the AP2 domains (the YRG element) is basic and hydrophilic. The carboxy terminal RAYD element contains a central region that, in almost all AP2 domains, is predicted to adopt the configuration of an alpha-helix of amphipathic character. Solution structure determination of one particular Arabidopsis AP2 domain protein revealed that the YRG element and the next 20 to 30 amino acids form three anti-parallel β-sheets that contact the DNA, whereas the RAYD element indeed adopts an α-helical structure as predicted (Allen et al., 1998).

As used herein an AP2 DNA-binding domain is defined as an amino acid sequence having at least 35% amino acid residue identity such as at least 45%, at least 55%, at least 65%, at least 75%, at least 85%, at least 95% or at least 98% residue identity to any of the AP2 domains described in Riechmann and Meyerowitz (1998). However, in the present context, any amino acid sequence having similarity to any of the two conserved segments, the YRG element and the RAYD element as defined by Riechmann and Meyerowitz (1998) are encompassed by the invention and considered as having an AP2 DNA-binding domain.

In another embodiment, the transcription factor as used herein includes at least two AP2 domains such as at least three AP2 domains.

In specific embodiments, the method is based on the use of a transcription factor coding sequence comprising at least one sequence coding for an AP2 domain selected from AP2 domains of SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7 or a sequence coding for an amino acid sequence selected from SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7.

It will be appreciated that in accordance with the invention, a functional AP2 domain class transcription factor can be derived from a source cell naturally containing such a factor. However, it is contemplated that a transcription factor that is functional in accordance with the invention can be constructed either by isolating one or more AP2 domains from a naturally occurring transcription factor and combining it with an amino acid sequence with which it is not naturally associated. Alternatively, an AP2 domain may be constructed

synthetically and combined with further amino acids to provide a transcription factor active polypeptide.

In accordance with the method of the invention, the nucleotide sequence may be a sequence coding for an amino acid sequence comprising an at least one AP2 domain having at least 40% homology, at least 50%, at least 60%, at least 70%, at least 80% or at least 90% homology to any of the at least one AP2 domains of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

As used herein, the term "homology" refers to homology at the amino acid level both in terms of amino acid identity and similarity. Similarity allows for "conservative variation", i.e. substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine with another hydrophobic residue, or the substitution of one polar residue with another polar residue, such as arginine with lysine, glutamic acid with aspartic acid, or glutamine with asparagine. In the present context, similarity is defined and determined using the present version of the BLAST program (Altschul et al, 1997).

In other specific embodiments, the nucleotide sequence coding for the transcription factor comprises at least one AP2 domain-coding sequence selected from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 or SEQ ID NO:4 or the nucleotide sequence coding for the at least one AP2 domain is a mutant, allele, derivative or variant of a nucleotide sequence coding for an AP2 domain of a sequence selected SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 or SEQ ID NO:4.

In the present context, the terms "mutant", "allele", "derivative" or "variant" may be by one or more of addition, insertion, deletion or substitution of one or more nucleotides in the nucleotide sequence, leading to the addition, insertion, deletion or substitution of one or more amino acids in the encoded polypeptide while substantially maintaining biological activity of the encoded polypeptide. Also included are changes in the nucleotide sequences which make no difference to the encoded amino acid sequence.

Variant, allelic, mutant or derivatised nucleotide sequences can e.g. be produced by standard DNA mutagenesis techniques or by chemically synthesising the variant DNA molecule. Such variants, alleles, mutants or derivatives do not change the reading frame of

the protein-coding region of the nucleotide sequence and encode a protein having no or only minor changes in its biological function.

In accordance with the method of the invention, the nucleotide sequence coding for a transcription factor, the expression of which is modulated, is either a homologous nucleotide sequence or a heterologous nucleotide sequence. As used herein, the term "homologous" refers generally to a nucleotide sequence present or originating in the same species and the term "heterologous" refers correspondingly to a sequence originating from a different species.

It will be appreciated that when the sequence is a homologous sequence, the expression of the sequence may be modulated while it is in its natural location in the cell, but it is also contemplated that a homologous coding nucleotide sequence may be translocated within the cell of origin. Whether homologous or heterologous, the transcription factor-encoding nucleotide sequence may be operably linked to a regulating nucleotide sequence with which it is not natively associated.

In one step of the above method, a nucleotide sequence coding for a transcription factor is inserted into the cell. The insertion can be carried out using any conventional methods of inserting nucleic acid into cells. Such transformation methods include as examples microprojectile bombardment, microinjection, electroporation, liposome mediated uptake and transformation by means of a disarmed Ti-plasmid vector carried by *Agrobacterium*. Numerous plant transformation vectors are available for plant transformation, and genes encoding transcription factors according to the invention can be used in conjunction with any such vectors. The selection of vectors for use will depend upon the preferred transformation technique and the target species for transformation. For certain target species, different antibiotic or herbicide selection markers may be preferred. Selection markers used routinely in transformation include the *nptII* gene which confers resistance to kanamycin and related antibiotics the *bar* gene which confers resistance to the herbicide phosphinothricin, the *hph* gene which confers resistance to the antibiotic hygromycin, the *dhfr* gene which confers resistance to methotrexate and the *Tdc* gene, which confers resistance to 4-methyltryptophan.

Also many vectors are available for plant transformation using *Agrobacterium tumefaciens*. These typically carry at least one T-DNA border sequence and include vectors such as

pBIN19, pCAMBIA, pMOG402 and pMOG22. Typical vectors contain DNA borders for *Agrobacterium*-mediated transformation, the RK2-derived tra function for mobilisation between *E. coli* and other hosts, and the OriT and OriV functions that are also from RK2. They usually contain a polylinker that is suitable for the cloning of plant expression cassettes containing their own regulatory signals.

Transformation without the use of *Agrobacterium tumefaciens* can also be performed which circumvents the requirement for T-DNA sequences in the chosen transformation vector, and consequently vectors lacking these sequences can be utilised in addition. The choice of vector depends largely on the preferred selection for the species being transformed.

In accordance with the invention, gene sequences intended for expression in transgenic plants can be assembled in expression cassettes behind a suitable promoter and upstream of a suitable transcription terminator. These plant expression cassettes can then by use of standard methods be transferred to the plant transformation vectors described above.

The selection of a promoter used in expression cassettes will determine the spatial and temporal expression pattern of the transgene in the transgenic plant. Selected promoters will express transgenes in specific cell types (such as leaf epidermal cells, mesophyll cells, root cortex cells) or in specific tissues or organs (roots, leaves or flowers, for example) and this selection will reflect the desired location of expression of the transgene. Alternatively, the selected promoter may drive expression of the gene under a light-induced or other temporally regulated promoter. A further alternative is that the selected promoter be chemically regulated. This provides the possibility of inducing expression of the transgene only when desired and caused by treatment with a chemical inducer.

A variety of transcriptional terminators are available for use in expression cassettes. They are responsible for the termination of transcription beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators are those which are known to function in plants and include the CAMV 35S terminator, the *tml* terminator, the nopaline synthase terminator, and the pea rbcs E9 and 3C terminators. They can be used in both monocotyledons and dicotyledons.

It is a further objective of the present invention to provide the means of enhancing biosynthesis of metabolites in a cell as defined above. Although it is contemplated that the

discoveries on which the present invention is based is applicable to any of these cell types and to any metabolite synth sised in such cells by genes the expression of which are regulated by transcription factors as defined herein, the method of the invention is presently preferred for the optimisation of biosynthesis of plant metabolites, in particular secondary metabolites including those mentioned above. However, It is also contemplated that the method of the invention can be used for the biosynthesis of novel and presently unknown metabolites, including novel plant secondary metabolites.

Thus, the secondary metabolites the biosynthesis of which can be enhanced using the method of the invention include alkaloid compounds including indole alkaloids such as terpenoid indole alkaloids (TIAs), phenolic compounds and terpenoid compounds. A review of secondary plant metabolites is included in e.g. Phytochemical Dictionary, A Handbook of Bioactive Compounds from Plants, edited by Harborne and Baxter, Taylor & Francis, London and Washington, DC. In useful embodiments, the secondary metabolites are alkaloids derived from compounds such as tryptophan (indole alkaloids including vincristine and vinblastine), tyrosine (isoquinoline alkaloids including berberine, berbamine and stephamine), lysine (quinolizidine alkaloids including lupamine and sparteine), ornithine (tropane alkaloids including scopolamine and atropine), nicotinic acid, anthranilic acid and acetate.

In accordance with the invention the transcription factor as used in the above method is capable of regulating the expression of at least one gene coding for a gene product involved in metabolite biosynthesis. Typically, such a gene product is a protein such as an enzyme. In a useful embodiments the enzyme is involved in an alkaloid biosynthetic pathway including an enzyme that is selected from tryptophan decarboxylase (TDC), geraniol-10-hydroxylase (G10H), strictosidine synthase (STR) or strictosidine glucosidase (SG).

As mentioned above, the present invention has made it possible to modulate in a cell the expression of one or more genes involved in the biosynthesis of a metabolite or a precursor for the metabolite. Although in presently preferred embodiments the expression of the gene or genes is enhanced with the objective of improving the yield of a desired metabolite, it is also contemplated that it may, under certain conditions, be advantageous to be able to reduce the expression of one or more biosynthetic genes. The enhancement, or alternatively, the reduction of the gene expression, relative to a cell into which a

nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor, is obtained by modifying the expression level of the nucleotide sequence coding for the transcription factor as defined herein.

In a specific embodiment, the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is enhanced by at least 10%, at least 25%, at least 50%, at least 75% or at least 100%, relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor. It will be understood that "enhancement of gene expression" is related to the increase in biosynthesis of the metabolite. It is also contemplated that the enhancement of gene expression can be higher such as by a factor of at least 1.5 including at least 2, at least 5 or at least 10.

The enhancement of the expression of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor can, in accordance with the invention, be achieved in either of two ways. One way is by inserting into the cell a nucleotide sequence coding for a transcription factor as defined herein and operably linking it to at least one expression regulating sequence, the cell comprising at least one target biosynthetic gene for the inserted transcription factor. An alternative way is by modifying the expression of a nucleotide sequence coding for such a transcription factor already present in the cell. In either of such ways, the modification results in that the transcription factor is produced in higher amounts than in a cell of a corresponding type that is not modified by insertion of a transcription factor-encoding sequence or by modifying the expression of such a sequence already present in the cell. It is evident that the expression of the biosynthetic gene is enhanced in a cell type not previously containing the transcription factor provided the cell contains a biosynthetic gene comprising a target sequence for the transcription factor. However, the expression level of an inserted transcription factor-encoding nucleotide sequence can also be enhanced by *in vivo* or *in vitro* modifications.

In accordance with the invention, enhancement of the expression of a transcription factor encoding sequence ("upregulation") can be obtained by any known technique leading to enhanced gene expression. As an example, the coding sequence can be linked to a stronger promoter or the native promoter of the sequence can be structurally modified so

as to improve its promoter strength. The expression may also be enhanced by modifying any other regulating sequence operably linked to the coding sequence e.g. by mutation. An alternative approach is to modify the nucleotide sequence of the transcription factor-encoding sequence.

A number of different sequences have been found to enhance gene expression from within the transcriptional unit and these sequences can thus be used in conjunction with the transcription factor-encoding nucleotide sequences according to the invention to increase their expression in transgenic plants. Various intron sequences have been shown to enhance expression, particularly in monocotyledonous cells. Intron sequences have been routinely incorporated into plant transformation vectors, typically within the non-translated leader.

A number of non-translated leader sequences derived from viruses are also known to enhance expression, and these are particularly effective in dicotyledonous cells. Specifically, leader sequences from Tobacco Mosaic Virus, Maize Chlorotic Mottle Virus and Alfalfa Mosaic Virus have been shown to be effective in enhancing expression.

Where the expression of the at least one gene involved in the metabolite biosynthesis is reduced ("down-regulated"), the reduction is typically by at least 10%, at least 25% or at least 100%, relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor. It is also contemplated that the reduction of gene expression can be higher such as by a factor of at least 1.5 including at least 2, at least 5 or at least 10.

Down-regulation of the expression of the transcription factor-encoding sequence can be obtained in several ways. Thus, as an example the expression is reduced by using anti-sense technology or "sense regulation" ("co-suppression"). In using anti-sense genes or partial gene sequences to down-regulate gene expression, a nucleotide sequence is placed under the control of a promoter in "reverse orientation" such that transcription yields RNA which is complementary to normal mRNA transcribed from the sense strand of the target gene. An alternative approach is to use a copy of all or part of the target gene inserted in sense, that is the same, orientation as the target gene, to achieve reduction in expression of the target gene by co-suppression.

In another aspect of the invention there is provided a method of modulating the stress resistance of a cell, the method comprising inserting into the cell a nucleotide sequence coding for a transcription factor, the nucleotide sequence is operably linked to at least one expression regulating sequence, and/or modifying the expression of a nucleotide sequence coding for such a transcription factor already present in the cell, the transcription factor is capable of regulating the expression of at least one gene coding for a gene product involved in the biosynthesis of said metabolite or precursor, and subjecting the cell to conditions where the inserted nucleotide sequence coding for a transcription factor is expressed or the expression of the nucleotide sequence already present in the cell is modulated, the expression of said nucleotide sequence resulting in a modified responsiveness of the cell towards exogenous stress conditions.

In the present context, the expression "exogenous stress conditions" should be understood as meaning any condition that results in that the growth and function of the cell is impaired relative to optimum conditions for growth and function. In relation to a plant, such stress conditions include attack of the plant by a pathogen, abiotic adverse environmental conditions like cold or heat and drought (osmotic stress), insufficient or excess supply of nutrient compounds, insufficient or excess levels of plant hormones, UV light, low light levels and suboptimal pH conditions.

In accordance with the invention the cell having modulated stress resistance may be any of the cell types as mentioned above in which at least one gene involved in the biosynthesis of a metabolite or a precursor for the metabolite is modulated as described above and to which there is referred. In particularly useful embodiments the cells having modulated stress resistance are plant cells of the above types and species including whole transgenic plants.

In accordance with the present method of modulating stress resistance, the expression of at least one biosynthetic gene is modulated by modifying the expression of any type of transcription factors that has a target sequence in the cell. Thus, in one embodiment, the nucleotide sequence coding for a transcription factor is a sequence coding for a transcription factor comprising at least one AP2 domain, including at least two AP2 domains or at least three AP2 domains of any of the above types and having any of the

above sequences including such sequences that are operably linked to a regulating nucleotide sequence with which it is not natively associated.

In accordance with the invention the plant cell which is modified to have enhanced stress resistance may be modulated in its expression of any of the above metabolites or gene products involved in metabolite biosynthesis.

Similar to what is described above in the context of the method of modulating in a cell the expression of one or more genes involved in the biosynthesis of a metabolite or precursor herefor, the present method includes that the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is enhanced or reduced relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor. Such an enhancement or reduction is achieved *mutatis mutandis* as it is described for the former method.

In a further aspect, the invention provides a recombinant cell having, relative to its parent cell, enhanced or reduced biosynthesis of a metabolite or a precursor therefor, and/or enhanced or reduced expression of a gene product involved in metabolite production, the cell comprising a nucleotide sequence coding for a transcription factor other than a transcription factor having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating the expression of at least one gene coding for a gene product involved in the metabolite production, said sequence is inserted into the cell and/or its expression is modified by operably linking it to a regulating sequence with which it is not natively associated. In useful embodiments such a cell is a plant cell of any of the types and species as mentioned above and modified as also described above. In other useful embodiments such a cell is a cell wherein at least one pathway leading to the biosynthesis of at least one precursor for the metabolite can be stimulated in the cell in the presence of an inducing agent. Such inducing agents include chemical compounds (e.g. an elicitor, jasmonate and a hormone) and physical factors (e.g. light, temperature, pH and shear stress).

In another aspect, the invention provides a method of producing a metabolite including a plant secondary metabolite, the method comprising providing a recombinant cell as described above, cultivating said cell under conditions where the nucleotide sequence

coding for the transcription factor regulating the expression of at least one gene coding for a gene product involved in the metabolite production is expressed, and recovering the metabolite. The cultivation of the recombinant cell of the invention is carried out using any conventional cultivation media and any conventional culturing methods for the particular type of cell. Such conventional culturing methods also includes methods where the cell is cultivated under conditions where at least one precursor for the metabolite is added in a form which can be assimilated by the cell. In useful embodiments such precursors may include compounds such as e.g. tryptophan, tyrosine, lysine, ornithine, nicotinic acid, anthranilic acid and acetate. Also included are culturing methods where the recombinant cell of the invention is further genetically modified in order to stimulate the production of a precursor for the metabolite, and methods wherein a compound which stimulates the production of a precursor for the metabolite is added to the cell in form which can be assimilated by the cell.

In a particular aspect of the invention there is provided a method of constructing a recombinant cell as described above, the method comprising the steps of (i) identifying in a source cell at least one nucleotide sequence coding for a transcription factor other than a transcription factor having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor(s) is/are capable of regulating in the source cell the expression of at least one gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, (ii) isolating said nucleotide sequence and (iii) inserting said isolated nucleotide sequence(s) into a host cell comprising a gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, the insertion of said nucleotide sequence(s) resulting in that the expression of the gene of the host cell is modulated relative to the expression level in a host cell not comprising the inserted nucleotide sequence(s). In useful embodiments, at least two isolated nucleotide sequences are inserted into the host cell.

A typical non-limiting example of how the method is carried out is described in the below examples.

In further aspects, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence coding for an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7 including such an isolated nucleic acid that comprises a nucleotide sequence which is selected from the group

consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4 and a mutant, allele or variant hereof and a polypeptide which includes an amino acid sequence having at least 40% homology including at least 50%, such as at least 60%, at least 70%, at least 80% and at least 90% to an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

There is also provided a polypeptide having the amino acid sequence shown in SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7.

The invention also relates to a vector comprising a nucleic acid molecule as defined above including such as a vector that comprises at least one further sequence coding for a transcription factor and to a host cell comprising such a vector. The host cell is selected from the group consisting of a bacterial cell, a fungal cell, a yeast cell, an animal cell, a unicellular eukaryotic cell, an algae cell and a plant cell.

As mentioned above, the invention also provides a method of isolating a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain. A typical example of a probe that is useful in such a method is a fragment of a nucleic acid as shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 or SEQ ID NO:4 and a mutant, allele or variant hereof.

The invention also provides a method of isolating a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain by using a nucleotide sequence as shown in SEQ ID NO:1, 2, 3 and 4 in a polymerase chain reaction.

In other aspects, there are provided an isolated nucleic acid having the nucleotide sequence as shown in SEQ ID NO: 8 and a mutant, allele or variant hereof, and a method using this nucleic acid sequence for isolating a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain.

The invention will now be further illustrated by the following non-limiting examples and the drawings wherein

Fig.1 shows that the RV region of the *Str*1 BH promoter is required for elicitor and jasmonate responsiveness. Two transformed cell lines for each construct were subjected to a control treatment (C), or induced with purified elicitor (PE) or methyl-jasmonate (MJ). Constructs were introduced in *C. roseus* cell lines MP183L and BIX via particle bombardment and *Agrobacterium,* respectively;

Fig. 2 illustrates that the RV region of the *Str*1 promoter is an autonomous elicitor and jasmonate responsive element. Two transformed cell lines for each construct were given a control treatment (C) or induced with purified elicitor (PE), or methyl-jasmonate (MJ). Constructs were introduced in *C. roseus* cell line MP183L via particle bombardment;

Fig. 3 shows that mutations M2, 3, 4, and 5 in the RV region abolish elicitor and jasmonate responsiveness of the *Str*1 BH promoter. Two transformed cell lines for each construct were given a control treatment (C) or induced with methyl-jasmonate (MJ). Constructs were introduced in *C. roseus* cell line BIX via *Agrobacterium*;
Fig. 4 shows that mutations M3, 4, and 5 in the RV region abolish elicitor and jasmonate responsiveness of the *Str*1 BH promoter and that mutant M2 weakens the responsiveness. Two transformed cell lines for each construct were subjected to a control treatment (C) or induced with purified elicitor (PE) or methyl-jasmonate (MJ). Constructs were introduced in *C. roseus* cell line MP183L via particle bombardment;

Fig. 5 shows *Orca-1* and *Orca-2* expression patterns in elicited *C. roseus* suspension cultures. *C. roseus* cells were treated with 50 μM methyl-jasmonate (MJ), 0.5% crude yeast extract elicitor (YE) or with an equivalent amount of partially purified elicitor (PE). Cells were harvested after the time points indicated in the figure. RNA was extracted and Northern blots were probed with DNA sequences specific for *Orca-1* or *Orca-2*;

Fig. 6 illustrates that T-DNA activation tagging of the Orca-3 gene activates expression of the TIA biosynthetic genes tryptophan decarboxylase (TDC), strictosidine synthase (STR) and strictosidine glucosidase (SG). Northern blot analysis of RNA extracted from line 46 and a 4-mT-resistant control cell line (C). Blots were probed with *Tdc*, *Str*1 and *Sg* (6A). RNAs in figure 6B were hybridised with the

rescued flanking plant DNA from line 46, containing the complete *Orca-3* open reading frame;

Fig. 7 shows *Orca3* expression patterns in elicited *C. roseus* suspension cultures. *C. roseus* cells were treated with 50 μM methyl-jasmonate (MJ), 0.5% crude yeast extract elicitor (YE) or with an equivalent amount of partially purified elicitor (PE). Cells were harvested after the time points indicated in the figure. RNA was extracted and Northern blots were probed with the rescued flanking plant DNA from line 46, containing the complete *Orca-3* open reading frame;

Fig. 8 illustrates that ORCA-1 and 2 are transcriptional activators of the *Str*1-promoter in *C. roseus* cells. GUS/CAT ratios with standard deviation were calculated from duplicate enzyme activity measurements of three bombarded samples arising from one batch of coated particles in a representative bombardment experiment. Tungsten particles were coated with a mixture of plasmids containing BH-GUS, empty pMOG184, or pMOG184 carrying *Orca*-1 or *Orca*-2 in sense or anti-sense orientation with respect to the CaMV 35S promoter. To correct for transformation efficiency, a plasmid carrying CaMV35S-CAT was added to all coatings;

Fig.9 illustrates that ORCA-1 and 2 transcriptionally activate the *Str*1-promoter in *C. roseus* cells in a sequence-specific manner. GUS/CAT ratios with standard deviation were calculated from duplicate enzyme activity measurements of three bombarded samples arising from one batch of coated particles in a representative bombardment experiment. Tungsten particles were coated with a mixture of GUS plasmids and pMOG184 derivatives. The GUS gene was driven by the *Str*1 BH promoter or BH mutant derivatives M1 through M6. pMOG184 plasmids were either empty (pM), or carried *Orca*-1 (O1) or *Orca*-2 (O2) in sense orientation with respect to the CaMV 35S promoter. To correct for transformation efficiency, a plasmid carrying CaMV35S-CAT was added to all coatings;

Fig. 10 shows that ORCA-3 activates expression of the *Str*1NH promoter (-202 to −1 with respect to ATG start codon). *C. roseus* cells were bombarded with tungsten particles coated with a mixture of plasmids containing *Str*1 NH-GUS, empty pMOG184, or pMOG184 carrying *Orca*-3 in sense orientation with respect to the

CaMV 35S promoter, and stained for GUS enzyme activity by addition of the substrate X-Gluc;

Fig. 11 shows that ORCA-3 activates expression of the Tdc -219 promoter (-219 to +86 with respect to ATG start codon). *C. roseus* cells were bombarded with tungsten particles coated with a mixture of plasmids containing Tdc -219-GUS, empty pMOG184, or pMOG184 carrying ORCA-3 in sense orientation with respect to the CaMV 35S promoter, and stained for GUS enzyme activity by addition of the substrate X-Gluc; and

Fig. 12 shows that ORCA-1, -2 and -3 activate expression of the jasmonate-responsive RV element from the *Str*1-promoter. *C. roseus* cells were bombarded with tungsten particles coated with a mixture of plasmids containing 4RV-GUS, empty pMOG184, or pMOG184 carrying *Orca*-1, -2 or -3 in sense orientation with respect to the CaMV 35S promoter, and stained for GUS enzyme activity by addition of the substrate X-Gluc.

EXAMPLES

The techniques applied in the below examples are all standard recombinant DNA and molecular cloning techniques well known in the art and are e.g. described by Sambrook et al. (1989) Molecular cloning: A laboratory manual, second edition, Cold Spring Harbor Laboratory Press and Ausubel et al. (1987) Current protocols in molecular biology, Greene Publishing Associates and Wiley Interscience, New York.

EXAMPLE 1

1.1 Identification of an elicitor- and jasmonate-responsive element in the *Str*1 promoter

The RV region of the promoter of the *str1* gene (DDBJ/EMBL/GenBank database accession number Y10182) of *Catharanthus roseus* (Madagascar periwinkle) coding for strictosidine synthase was, as described below, via promoter analysis including detailed scanning mutagenesis, identified as an elicitor- and jasmonate-responsive element.

The BglII/HincII fragment (BH) of the *Str*1 promoter and a BH fragment in which the RsaI/AvaII fragment (RV) was deleted (BH^{-RV}) were introduced as BglII/SalI fragments into the GusSH vector consisting of pBluescript SK+ carrying the β-glucuronidase reporter gene. These *Str*1 promoter derivatives fused to the *gusA* gene were introduced into two different *C. roseus* cell lines. Cell line MP183L was transformed via particle bombardment, whereas cell line BIX was transformed via an improved *Agrobacterium* strain (Memelink, unpublished results, see example 4.1). For *Agrobacterium*-mediated transformation, the GusSH derivatives were introduced into the binary vector pMOG22lambdaCAT, a derivative of pMOGlambdaCAT containing the hygromycin resistance gene instead of the kanamycin resistance gene as a plant selectable marker. Northern blot analysis was performed for two independently transformed cell lines for each construct, each estimated to consist of a population of at least 100 to 1000 independent transformants. The BH fragment of the *Str*1 promoter conferred low basal level expression onto the *gusA* gene in the control (0.05% dimethylsulfoxide (DMSO)) treated cells (Fig.1). Incubation of MP183L cells transformed with BHGusSH with partially purified yeast extract elicitor (PE; Menke et al., 1999) and methyl-jasmonate (MJ) for 6 h induced the expression of the *gusA* gene several fold (Fig. 1 upper panel). These conditions similarly induced the endogenous *Str*1 gene as described previously (Menke et al.,1999). The BH^{-RV} fragment did not confer elicitor- and MJ-responsive gene expression onto the *gusA* reporter gene in cell line MP183L (Fig.1 upper panel). Similar results were obtained with BIX cell lines transformed with these *Str*1-promoter derivatives fused to gusA. BIX cells carrying the BH-*gusA* fusion had low basal *gusA* mRNA levels, but incubation for 6 h with MJ induced high *gusA* gene expression (Fig.1 lower panel). BIX cells carrying BH-RV fused to *gusA* did not have detectable basal expression levels and could not be induced by MJ (Fig.1 lower panel). This loss of function experiment shows that the RV region of the *Str*1 promoter is required for elicitor- and jasmonate-responsive gene expression.

In a gain of function experiment it was tested whether the RV fragment of the *Str*1 promoter was able to confer elicitor-and JA-responsive gene expression onto a minimal promoter-*gusA* fusion. To this end tetramers of the RV region were constructed and fused to a Caulifower Mosaic Virus 35S promoter truncated to -47 and fused to *gusA* in the vector GusSH-47. The RV region was cloned as a RsaI/AvaII fragment into the SmaI/EcoRV sites of pIC20H, and tetramerized using the enzymes BamHI/BglII. The tetramer was cloned as a BamHI/BglII fragment into the BamHI site of GusSH-47. *C. roseus* cell line MP183L was

transformed with this construct via particle bombardment. Control cell lines were generated by introducing GusSH-47 or 6Tcyt-GusSH-47 in two independently transformed cell lines for each construct. The 6Tcyt construct contains 6 head-to-tail copies of the cyt-1 cis-acting element from the T-DNA T-CYT gene which is not elicitor- or jasmonate-responsive. Northern blot analysis was performed for two independently transformed cell lines, each estimated to consist of a population of at least 100 to 1000 independent transformants. Incubation of the 4RV-47 cell lines with PE or MJ for 6 h resulted in induction of *gusA* gene expression as compared to the expression of the control (0.05% DMSO) incubated cells (Fig.2). The same treatment of cell lines carrying the -47 construct did not result in measurable *gusA* mRNA levels in the control or induced samples (Fig.2). In the cell lines carrying the 6Tcyt-GusSH-47 construct, high basal levels of *gusA* mRNA were found (Fig.2). Treatment of these cell lines with PE or MJ for 6 h did not induce expression (Fig.2). These results show that the RV region is sufficient to confer elicitor- and jasmonate-responsive gene expression onto an otherwise inactive minimal promoter-*gusA* construct. The results with the 6Tcyt lines show that the RV region does not potentiate a silent elicitor-responsive element within the CaMV -47 region or elsewhere on the construct.

Detailed block scanning mutagenesis of the RV region in the context of the BH fragment of the *Str*1 promoter was performed to pinpoint the exact location of the elicitor- and jasmonate-responsive-sequence within the RV fragment. Mutations were introduced by PCR with primers having a 6 bp mismatch combined with M13-reverse primer using the XhoI/HincII fragment of the *Str*1 promoter cloned in pIC20H as a template. The primers RVM1-RVM8 are disclosed in the below Table 1.1

Table 1.1 Primers used in detailed block scanning mutagenesis of the RV region

RVM1	5'-CCACGTGGTTGTAGTCTCTTAGACC-3'
RVM2	5'-GGTACATCAGAGAATGACCGCCTTC-3'
RVM3	5'-CACTCTTACTGGCGCTTCTTTGAAAG-3'
RVM4	5'-AGACCGCGAAGAATGAAAGTG-3'
RVM5	5'-AGACCGCCTTCTTACTTTCTGATTTCCCC-3'
RVM6	5'-TTGAAAGACTAAACCCTTGGAC-3'
RVM7	5'-GAAAGTGATTTGGGAACGACCTTG-3'
RVM8	5'-CCCTTGCTGGAAGTTTGGTGAG-3'

PCR fragments obtained were used as primers in PCR reactions together with M13-forward primer on the XhoI/HincII fragment of the *Str*1 promoter cloned in pIC20H. PCR fragments obtained were reamplified using M13-reverse and M13-forward primers.

Amplified PCR products were digested with BglII/SalI and cloned into GusSH digested with BamHI/SalI. In total 8 mutated versions of the RV fragment were made, covering the entire fragment from the RsaI site to the AvaII site. In each mutant 6 adjacent nucleotides (nt) were changed into their complementary nt, i.e. changing A to T and G to C. The 8 mutated versions of the BH fragment fused to *gusA* were introduced into *C. roseus* cell lines MP183L and BIX as described above. Northern blot analysis was performed for two independently transformed cell lines for each mutant construct, that were incubated for 6 h with DMSO or MJ for *C. roseus* BIX cell lines (Fig.3) and with DMSO, PE or MJ for *C. roseus* MP183L cell lines (Fig.4). By comparing basal and MJ-induced *gusA* gene expression for all the mutant BIX lines to the expression conferred by wildtype BH fragment it is clear that mutations 1, 6, 7 and 8 conferred a wildtype gene expression pattern (Fig.3). In mutant 2 the expression was only minimally inducible. In mutants 3, 4 and 5 expression was uninducible, and similar to the expression conferred by the BH-RV fragment (Fig.3). The experiment with the MP183L lines carrying the mutated BH fragment gave similar results (Fig.4). Mutants 1,2 and 6 showed expression patterns similar to the wildtype BH fragment. Mutants 3, 4 and 5 are not consistently inducible by PE or MJ and gave similar expression patterns to the BH-RV fragment.

In another experimental series, the effect of the block mutations was studied in constructs consisting of tetramers of RV derivatives fused to GusSH-47. Tetramers of RV mutants M2, M3, M4, M5 and M6 were constructed as described above for RV wildtype fragment, and fused to a Caulifower Mosaic Virus 35S promoter truncated to -47 fused to gusA in the vector GusSH-47. The constructs were introduced in *C. roseus* cell line MP183L via particle bombardment. Two independent transformed cell lines were generated for each construct. Incubation of the cell lines carrying RV wildtype or mutant M6 tetramer constructs with PE or MJ for 6 hours resulted in induction of gusA gene expression as compared to the expression of the control (0.05% DMSO) incubated cells (results not shown). The same treatment of cell lines carrying the mutant M2, M3, M4 or M5 tetramer constructs did not result in measurable gusA mRNA levels in the control or induced samples (results not shown).

These results show that the elicitor- and MJ-responsive element is contained within the 24 bp sequence that is covered by mutations 2 to 5, with mutations 3, 4 and 5 having the most dramatic negative effect on the inducibility of the BH or the RV fragment of the *Str1* promoter.

EXAMPLE 2

Cloning of ORCA-1 and 2

The RV region was, as described in Example 1, identified as an elicitor- and jasmonate-responsive element in the *Str*1 promoter. Use of the RV element as bait in a one-hybrid transcription factor screening resulted in the isolation of 17 cDNA clones that corresponded to two different mRNA species, encoding proteins belonging to the APETALA2 domain class of transcription factors. They were named ORCA-1 and 2 (for Octadecanoid-Responsive Catharanthus AP2).

The RV region of the *Str*1 promoter (DDBJ/EMBL/GenBank database accession number Y10182) was cloned as a RsaI/AvaII fragment into the SmaI/EcoRV sites of pIC20H and tetramerized using the enzymes BamHI/BglII. The tetramer was cloned as a BamHI/BglII fragment into the BamHI site of pHIS3NX. The tetramer-HIS3 gene fusion was transferred as a NotI/XbaI fragment to the yeast integration plasmid pINT1.

The resulting plasmid was linearized with NcoI and introduced into yeast strain Y187. Recombinants were selected on YPD medium containing 150 µg/ml G418, and the occurrence of a single recombination event between the pINT1 derivative and the chromosomal PDC6 locus was verified via Southern blotting. Yeast strain Y187-4RV was then used in a one-hybrid screen for DNA-binding proteins with a cDNA library of *Catharanthus roseus* cloned in the plasmid pACTII. To construct the library, cDNA was synthesised with a Stratagene kit on 2 different poly(A) RNA preparations mixed in a 1:1 ratio isolated from cell suspensions that were elicited with 0.05 % yeast extract for 1 and 4 h, respectively. The cDNA was cloned in the EcoRI/XhoI sites of lambdaACTII. The amplified lambda library, consisting of 3.5×10^6 independent primary transformants was converted to a pACTII plasmid library via Cre-lox-mediated in vivo excision in Cre-expressing in *E. coli* strain BNN132. Four million Y187-4RV transformants were screened in a one-hybrid screening. Plasmids isolated from 364 colonies obtained on SD medium without histidine and leucine were retransformed to Y187-4RV, and two Y187 derivatives carrying tetramers of RV mutant fragments RVM4 and RVM5 fused to the HIS3 gene. Seventeen plasmids showed strongly reduced growth in Y187-4RVM4 compared to Y187-4RV and Y187-4RVM5, whereas all other plasmids did not give rise to significant

differences in growth between the three yeast strains. Partial sequencing of these plasmids revealed that they belonged to one of only two classes. The ORCA-1 class consisted of 3 plasmids carrying cDNA sequences with identity to the sequences in SEQ ID Nos 1 and 2 . SEQ ID No:1 is derived from plasmid RV23, whereas SEQ ID No:2 is derived from plasmid RV124. Both sequences are identical throughout the main open reading frame, but diverge at their 3' ends. Obviously, they are derived from the same gene, and the differences could have arisen by alternative splicing of the primary transcript. They contain the complete coding region, since SEQ ID Nos. 1 and 2 contain an in frame stop codon preceding the first start codon. The ORCA-2 class consists of 14 plasmids carrying cDNA sequences with identity to the sequence in SEQ ID No:3. All ORCA-2 cDNA sequences appeared to be partial. The missing portion was isolated via PCR with the primers 4TH (5'-CCCCACCAAACCCAAAAAAAG-3') and RV117 (5'-CCATATCCTCGATCCTTTTCTC-3') using the pACTII cDNA library as a template. A prominent 0.6 kbp PCR band was digested with EcoRI and BamHI, and cloned in pBluescript II SK+. A complete clone was constructed in pBluescript II SK+ by fusion of the PCR fragment with the cDNA fragment from plasmid RV210 using the unique BamHI site.

EXAMPLE 3

Expression of ORCA-1 and -2

To analyse the expression of the Orca-1 and Orca-2 genes suspension-cultured *C. roseus* cells were incubated with MJ (50 µM) or PE (0.05% YE equivalent). After incubation periods ranging from 1 to 28 h the cells were harvested and immediately frozen in liquid nitrogen. Total RNA was isolated and Northern blots were probed with a 700 bp EcoRI fragment from SEQ ID NO:1, which is specific for Orca-1, or a 450 bp EcoRI/KpnI fragment from SEQ ID NO:3, specific for Orca-2. For Orca-1 the control sample, incubated with DMSO for 6 h, had high levels of Orca-1 mRNA, but even higher levels were induced by MJ after 4 h (Fig.5). MJ-induced gene expression was clear for Orca-2 with high levels of expression at 2 h as compared to the control. The level of mRNA remained high up to 28 h (Fig.5). Both Orca-1 and Orca-2 were rapidly induced by PE, with an increase of mRNA levels within 1 h for Orca-2 and within 2 h for Orca-1 (Fig.5). The induction of Orca-1 and Orca-2 gene expression peaked at 2 to 3 h after which the amount of both mRNAs declined. The same induction patterns were found for crude yeast extract (YE) (Fig.5).

These results show that both Orca-1 and Orca-2 are inducible by MJ and PE, with Orca-2 being induced most rapidly.

EXAMPLE 4

Cloning of ORCA-3

T-DNA activation tagging has been used as a reverse genetic method to isolate genes involved in plant growth and development. Genes involved in responses to ABA and cytokinin have been identified using this technique. A T-DNA carrying a strong transcriptional enhancer is integrated into the plant genome by *Agrobacterium tumefaciens*-mediated transformation. Expression of genes adjacent to the integration site will become deregulated, which can result in a dominant mutation.

To isolate dominant mutants with increased expression of TIA (terpenoid indole alkaloid) biosynthetic genes, the *Tdc* gene was used as selectable marker. TDC is involved in the conversion of L-tryptophan into tryptamine, one of the first steps in TIA biosynthesis. TDC is able to use substrates other than L-tryptophan, like 4-methyltryptophan (4-mT). This compound is toxic to plant cells, and is converted by TDC in to the non-toxic 4-methyltryptamine. We used 4-mT to select for *C. roseus* cell lines in which the activation tag causes an increase in *Tdc* expression. The selection for cell lines with increased *Tdc* expression is described below

4.1 Plasmids and *Agrobacterium tumefaciens* strain

Plant cell transformations were done by applying the ternary vector system (van der Fits et al., manuscript in preparation). A KpnI fragment carrying a constitutive *virG* mutant gene was cloned onto plasmid pBBR1MCS and this plasmid was used as ternary vector.

A brief outline of the cloning of the binary vector Tag-2B4A1 follows: The 35S promoter derivative 2B4A1, containing multiple copies of the B and A1 enhancers from the CaMV 35S promoter, was cloned on a pUC21 plasmid containing the *Hpt* gene driven by a deletion derivative of the FMV 34S promoter (34S-*Hpt*) and the *gusA* reporter gene. This complete plasmid (including pUC21 sequences) was digested at a unique BglII site and cloned into the binary vector pSDM14 cut with BamHI. The 35S promoter is now reading towards the right border sequence.

Both plasmids (binary and ternary vectors) were introduced into *Agrobacterium tumefaciens* strain LBA4404 via triparental mating using pRK2013 as a helper strain.

4.2 Transformation of Catharanthus roseus suspension cultured cells

Agrobacterium tumefaciens strains were grown for 3 days at 28°C on solid AB medium containing 20 µg/ml rifampicin, 100 µg/ml kanamycin and 75 µg/ml chloramphenicol. From these bacteria a liquid overnight culture was grown in the same medium. Cells were harvested by centrifugation and resuspended at a OD_{600} of 1 in LS medium of pH 5.2 supplemented with 2 mg/l NAA, 0.2 mg/l kinetin, 3% sucrose, 1% glucose and 100 µM acetosyringone (hereafter called cocultivation medium).

C. roseus cells (cell line BIX) are grown in LS medium supplemented with 3% sucrose, 0.2 mg/l kinetin and 2 mg/l NAA (hereafter called LS13) at 25°C (16/8 h light/dark cycle) and subcultured every 7 days by transferring 7.5 ml of cells in 50 ml medium. For transformation 7 days old cells were used. The culture medium was replaced by cocultivation medium and the cell density was adjusted to 50% sedimentation volume. On a petridish (diam. 80 mm) containing 25 ml of cocultivation medium solidified with 0.7% plant tissue culture agar, 7.5 ml of this *C. roseus* suspension was mixed with 750 µl *Agrobacterium tumefaciens* culture. Dishes were taped using leucopor tape and put at 25°C in the dark for 3 days. After the cocultivation the cells were washed and plated on a paper filter disc (diam. 75 mm, Whatmann, No.4). These filters were placed in petridishes containing 25 ml LS13 with 50 µg/ml hygromycin-B, 400 µg/ml cefotaxim, 100 µg/ml vancomycin and 0.4 mM 4-methyl-tryptophan. Filters were transferred onto fresh medium weekly.

4.3 Isolation of a 4-mT resistant *C. roseus* line with increased TIA biosynthetic gene expression

An estimated number of 400.000 to 500.000 stable, independent *C. roseus* transformants were generated by Agrobacterium tumefaciens-mediated transformation with the tagging construct Tag-2B4A1. Transformants were placed on selective medium containing 50 µg/ml hygromycin-B, 400 µg/ml cefotaxim, 100 µg/ml vancomycin and 0.4 mM 4-mT. After several weeks of selection 281 independent 4-mT-resistant calli were obtained. These calli

were transferred into liquid medium containing 50 µg/ml hygromycin-B and 0.2 mM 4-mT. From these suspension cultures RNA was extracted. Northern blot analysis showed that several mutants had an increased *Tdc* mRNA level (data not shown), which in all likelihood is causing the 4-mT resistance. One these mutants (line 46) is described in more detail below.

For line 46 Northern blot analysis was performed using other genes from the TIA biosynthetic pathway. RNA extraction and Northern blot analysis were performed. Ten µg RNA samples were loaded onto the gels. As 32P-labeled probes, cDNAs encoding TDC, STR (Pasquali et al., 1992), and SG (Geerlings, Memelink, van der Heijden and Verpoorte, unpublished results) were used.

Besides *Tdc*, also *Str*1 and *Sg* mRNA levels were increased significantly in the tagged cell line 46, when compared to a 4-mT resistant control cell line (number 38) (Fig.6A). Thus in line 46, the T-DNA tag appears to have activated a central regulator of TIA biosynthetic gene expression, thereby inducing the expression of at least three genes, and possibly the complete TIA biosynthetic pathway. Chromosomal DNA was extracted from 50 gram fresh weight of cells from suspension line 46 and isolated by CsCl-Ethidiumbromide ultracentrifigation. Southern blots were made and the *gusA* gene was used as 32P-labeled probe isolated from GusSH. Southern blot analysis of line 46 showed that only one T-DNA copy was integrated into the plant genome (data not shown).

For plasmid rescue, 10 µg genomic DNA of line 46 was digested with XbaI (a unique site in tagging construct Tag-2B4A1), ligated and transformed to electrocompetent cells of E.coli strain NM554.

Plasmid analysis of the obtained E.coli transformants identified the rescued plasmid 46X8, which carried 1.6 kb of plant DNA immediately flanking the right border of the tagging construct. Using the rescued sequence as a probe on a Northern blot of RNA isolated from line 46 and control line 38, one hybridising mRNA species was observed (Fig.6B). This transcript was absent in control cell line 38, whereas it was present at a high level in line 46.

Sequence analysis of plasmid 46X8 revealed an open reading frame (ORF), located at approximately 600 bp downstream of the start of the rescued plant sequence. Homology

searches in databases revealed one AP2 domain present in the predicted amino acid sequence of this ORF.

AP2 domain containing proteins are a class of transcription factors exclusively found in plants. The encoded protein was called ORCA-3 (for octadecanoid-responsive Catharanthus AP2 domain). Several AP2 domain transcription factors, all containing one conserved AP2 domain, have been shown to participate in stress signalling (Pti, EREBP, CBF1, DREB). CBF1 has been demonstrated to act as central regulator of cold-induced gene expression in Arabidopsis thaliana. In analogy, the tagged gene described here, appears to encode a central regulator of TIA biosynthetic gene expression.

The sequence rescued on plasmid 46X8 is of genomic origin, and may be interrupted by intron sequences. To isolate a cDNA copy of the mRNA transcribed from the tagged gene, a PCR approach was followed. Primer sequences were based on DNA sequence information from the rescued plasmid. Primer OR1 (5'-GAATTCATATGGCGGAAAGCTGTCAGGAGGATTC) was combined with the vector T3 primer and OR4 (5'-CGACGTCGTAGAAGGCTCCGCAGGG) with the vector M13-40 primer in PCR reactions using lambda DNA prepared from a *C. roseus* root cDNA library (Pasquali et al., 1992) as a template to isolate the 3' respectively 5' parts of the ORCA3 cDNA. Based on DNA sequence data from these partial cDNA clones, primers OR5 (5' AGATCTCATAGTTCCGAAGAAATC-ATTTCCGTCTCAG) and OR6 (5' AGACTCGTGAACTTTTTTGGATATAAAATTTTGTAC-ATTCC) were developed, which were used in a PCR on the rescued 46X8 plasmid to obtain the complete ORCA-3 open reading frame. The cDNA and amino acid sequences are provided in SEQ ID NO:4 and SEQ ID NO:7, respectively. Homology with other AP2 domain proteins, including the Catharanthus proteins ORCA-1 and ORCA-2 described above, is restricted to the AP2 domain. Sequences outside this domain are not conserved.

EXAMPLE 5

The tagged ORCA-3 gene is induced by elicitor and methyl-jasmonate

Previous research showed that the octadecanoid pathway, leading to the synthesis of jasmonate, is involved in elicitor-induced expression of TIA biosynthetic genes (Menke et

al., 1999). Treatment of *C. roseus* cells with methyl-jasmonate (MJ) thus resulted in a transient increase of both *Tdc* and *Str*1 mRNA levels.

The T-DNA tagging experiment in Example 4.3 showed that the *Tdc*, *Str*1 and *Sg* genes are regulated by ORCA-3. Based on the fact that two other *Catharanthus* AP2 domain proteins, ORCA-1 and 2, are involved in the elicitor- and jasmonate-responsive expression of the *Str*1 gene (Example 2), it was speculated that ORCA-3 may be a third elicitor- and jasmonate-responsive transcription factor.

Northern blots were loaded with RNA from *C. roseus* cells treated for various time points with MJ, crude yeast extract elicitor (YE) and with a partly purified elicitor from yeast extract (PE). Subsequently, blots were probed with plasmid 46X8, containing the tagged gene. Treatment of cells with MJ resulted in a rapid increase in expression of the tagged gene (Fig.7). This increase was already observed within 2 h after addition of MJ and sustained for at least 28 h. The induction of *Tdc* and *Str*1 expression was slower (maximal at 4 h after treatment) and was also maintained for at least 28 h (Menke et al., 1999). Because of the induction by a product from the octadecanoid pathway, the tagged gene was named *Orca-3* (Octadecanoid Responsive Catharanthus AP2-domain protein). Incubation of cells with crude (YE) and partly purified (PE) elicitor (Fig.7) resulted in a rapid and transient increase of *Orca-3* mRNA levels, although the expression level is low relative to MJ-induced levels. The maximal induction of gene expression is at approximately 2 h after addition of the elicitor. The maximal induction of *Tdc* and *Str*1 expression by the elicitors occurs later (maximal induction after 6-8 h; Menke et al., 1999).

The Northern blot analysis shows that *Orca-3* expression is elicitor- and jasmonate-responsive and precedes the induction of the TIA biosynthetic genes, indicating that ORCA-3 is an elicitor- and jasmonate-responsive transcription factor that coordinately controls TIA biosynthetic gene expression.

EXAMPLE 6

Binding site of ORCA-1 and 2 within the RV region of Str promoter

To identify the ORCA binding site within the RV region, Y187 yeast derivatives that contained tetramers of the RV wildtype fragment and mutants M1 through M6 (see below

and Table 1.1 for mutations) fused to the HIS3 gene were constructed via single cross-over in the PDC6 locus as described above.

The sequence of the RV region (SEQ ID NO:8) of the *Str*1 promoter is shown below. Mutated blocks are underlined or indicated in bold face. Nucleotides within each block were changed into their complementary nucleotides.

```
                M2              M4      M6
GTACATCACTCTTAGACCGCCTTCTTTGAAAGTGATTTCCCTTGGACC
     M1             M3              M5          M7
```

The yeast strains were then separately transformed with pACTII derivatives RV124 (encoding ORCA-1 fused to the GAL4 activation domain) and RV210 (encoding part of ORCA-2 including the AP2 domain fused to the GAL4 activation domain). Transformants were plated on SD minimal medium without leucine but containing histidine. Colonies obtained after 5 days were streaked on SD plates without histidine and leucine, and growth was recorded after 5 days (Table 6.1). Growth indicates that ORCA proteins bound the RV derivative and activated HIS3 gene transcription, whereas lack of growth indicates that ORCA proteins failed to bind the RV derivative. The results indicate that ORCA-1 and 2 failed to confer growth in yeast containing mutants M3 and M4 fused to the HIS3 gene. In these mutants, blocks of 6 nucleotides are changed in their complementary nucleotides within the sequence GACCGCCTTCTT. This sequence is similar to the binding site identified for tobacco EREBP proteins (Ohme-Takagi and Shinshi, 1995) and tomato Pti proteins (Zhou et al., 1997) within PR gene promoters and for Arabidopsis CBF and DREB proteins within COR gene promoters (Stockinger et al., 1997).

TABLE 6.1

Growth of yeast Y187 derivatives containing different RV mutants fused to the HIS3 gene transformed with ORCA-1 or ORCA-2 expressing pACTII plasmids.

RV derivative	ORCA-1 (RV124)	ORCA-2 (RV210)
RV	+	+
M2	+	+
M3	-	-
M4	-	-
M5	+	+
M6	+	+

EXAMPLE 7

ORCA-1 and 2 are transcriptional activators of the *Str*1-promoter in *C. roseus* cells

To investigate whether ORCA-1 and 2 can act as transcriptional activators interacting with the *Str*1-promoter, *C. roseus* cells were transiently cotransformed with a *Str*1-promoter-*gusA* fusion (BH-GusSH) construct , an overexpression vector (pMOG184) carrying *Orca1* or *Orca2* cDNAs fused in sense or antisense orientation to the CaMV 35S promoter and a overexpression vector carrying the *chloramphenical acetyl transferase* gene (*cat*). As a control, cotransformation of BH-GusSH with an empty overexpression vector was done. The CaMV 35S-cat construct served as an internal control for transformation efficiency. Cells were transformed through particle bombardment as described before, using the three constructs in a ratio of 1:1:3 (gus:cat:orca). Twenty four hours after transformation, cells were harvested and frozen in liquid nitrogen. GUS and CAT activity assays were performed and the *Gus* reporter gene expression was expressed as a GUS/CAT ratio to correct for transformation efficiency. Expression of *Orca-1* in sense orientation resulted in a higher GUS activity (gus/cat ratio, 1.6 fold induction) as compared to the control (BHGusSH +pMOG 184) (Fig. 8). Expression of *Orca-1* in antisense orientation did not alter the GUS activity level (Fig. 8). Expression of *Orca-2* in sense orientation resulted in a 3.3 fold induction of GUS activity, whereas expression of *Orca-2* in antisense orientation had little or no effect on GUS activity as compared to the control transformed cells (BHGusSH +pMOG 184) (Fig. 8).These results show that transient overexpression of either ORCA-1 or 2 (in sense orientation) enhances the expression of the gusA reporter gene driven by the Str-promoter fragment BH.

To show that ORCA-1 and 2 enhance the expression driven by the *Str*1-promoter in a sequence-specific manner, co-transformations were done using wild-type and mutant versions of the *Str*1-promoter (fused to gusA). ORCA-1 enhanced the expression of the gus reporter gene to some degree through interaction with wild-type and mutant versions 1, 2, 5 and 6 of the BH fragment of the *Str*1-promoter, as compared to the control pMOG184) (Fig. 9). ORCA-1 did not enhance the expression of reporter constructs driven by mutant versions 3 and 4 of the *Str*1-promoter fragment BH (Fig. 9). ORCA-2 strongly enhanced the expression of gus reporter gene through interaction with wild type and mutant version 1, and 6 of the BH fragment of the *Str*1-promoter (Fig. 9), as compared to the control (pMOG184). ORCA-2 enhanced gus gene expression weakly for mutants 2 and 5 (Fig. 9). ORCA-2 did not enhance the expression of reporter constructs driven by mutant versions 3 and 4 of the *Str*-promoter fragment BH (Fig. 9). These results show that the activation of gene expression by ORCA-1 and 2 is sequence-specific and occurs through direct interaction with the *Str*1-promoter fragment RV. The sequence requirement for ORCA activation in plant cells is slightly different from that in yeast cells (see example 6). In yeast, only mutations M3 and M4 affect the ability of ORCA's to activate gene expression, whereas in plant cells mutations M2 and M5 additionally affect gene expression to some degree. These differences may reflect differences in binding in the two cell types, for instance due to interaction with different endogenous proteins. On the other hand, the transient co-bombardment data closely correlate with the results of the stable transformation (Example 1), indicating that indeed lack of ORCA binding is the reason for lack of expression of mutants M2, 3, 4 and 5 in stably transformed lines.

EXAMPLE 8

ORCA-3 is a transcriptional activator of the *Tdc* and *Str*1 promoters in *C. roseus* cells

To investigate whether ORCA-3 activates expression from the *Tdc*- and *Str*1-promoters, a particle bombardment assay was performed. *C. roseus* cells were co-bombarded with the *gusA*-reporter gene with or without an overexpression vector containing the *Orca-3* cDNA in the sense orientation. The *gusA*-reporter gene was driven by a fragment of the *Tdc*-promoter (-219 to +86 as referred to the translational start site) or the *Str*1-promoter (-202 to -1 as referred to the translational start site).

24 h after bombardment the cells were stained for GUS activity. A clear increase in the number and intensity of blue spots was observed when the *gusA*-reporter gene was co-bombarded with 35S-*Orca3*, when compared to a negative control in which the cells were co-bombarded with the empty overexpression vector. This enhancement was observed both with the *Tdc*-promoter fragment (Fig. 11) as well as with the *Str*1-promoter fragment (Fig. 10) driving reporter gene expression. As a control, GUS activity from an unrelated promoter (6Tcyt) was not increased after co-bombardment with *Orca-3* (data not shown). Thus in this transient transformation assay, *Orca-3* activates gene expression from the *Tdc*- and *Str*1-promoter in a sequence-specific manner.

EXAMPLE 9

ORCA-1, -2 and -3 activate expression of the jasmonate-responsive RV element from the *Str*1-promoter

The transient co-bombardment assay, as described above, was used to investigate whether the ORCA proteins activate gene expression of the jasmonate-responsive RV element from the *Str*1-promoter. *C. roseus* cells were bombarded with a *gusA*-reporter gene driven by a tetramer of the RV element (4RV) in combination with the empty overexpression plasmid, 35S-*Orca1*, 35S-*Orca2* or 35S-*Orca3*. The number and intensity of blue spots was increased clearly upon co-bombardment with one of the three *Orca* genes, when compared to control cells, which were co-bombarded with the empty vector (Fig. 12). The strongest activation of *gusA* expression was observed when *Orca-2* or *Orca-3* were used. However, to make statements about the relative strengths of activation of the 3 different ORCAs, quantitative analysis is necessary, thereby correcting for differences in particle coating and bombardment efficiencies. In addition, the differences in activation may be due to differences in expression levels of the three ORCA proteins, instead of differences in activation strength.

EXAMPLE 10

ORCA-3 does not activate gene expression of RV mutants M2 and M3, and weakly for mutants M4 and M5

35S-*Orca3* was co-bombarded with the 4RV mutant series (described in Example 1) to determine which sequences in the RV element were important for the activation by *Orca3*. After co-bombardment cells were stained for GUS activity and blue staining was quantified visually by estimating the number and intensity of blue spots. The results are shown in the below Table 10.1.

TABLE 10.1

Relative GUS activities as determined visually

Cells bombarded with	Relative GUS staining		
4RV + pMOG	2	2	2
4RV + 35S-*Orca3*	4	4	4
4M2 + pMOG	1	1	1
4M2 + 35S-*Orca3*	1.5	1	1
4M3 + pMOG	0.5	0.5	0.5
4M3 + 35S-*Orca3*	0.5	0.5	1
4M4 + pMOG	0.5	0.5	0.5
4M4 + 35S-*Orca3*	1	1	1
4M5 + pMOG	0.5	0.5	0.5
4M5 + 35S-*Orca3*	1.5	1.5	2
4M6 + pMOG	3	2.5	2.5
4M6 + 35S-*Orca3*	4	4	4

The above Table 10.1 shows the relative GUS activities as determined visually. The *C. roseus* cells were co-bombarded in triplicate with the *gusA* reporter gene driven by 4RV or a mutant of this element together with an overexpression plasmid carrying the *Orca3* cDNA (35S-*Orca3*) or the empty vector (pMOG). Cells were stained for GUS activity. The numbers in Table 10.1 are an estimation of the intensity and number of blue spots. Values range from 0 (no blue spots) until 4 (many intense blue spots).

No activation of GUS activity was observed when the *gusA* reporter gene was driven by a tetramer of RV mutant M2 or M3. For mutants M4 and M5, the basal GUS activity was lower, when compared to the wild-type. However, ORCA-3 was able to activate gene expression of these mutants. Expression of mutant M6 resembled the wild-type, both in basal GUS activity levels, as well in inducibility by ORCA-3. Thus ORCA-3 is able to activate gene expression from the RV element and sequences that are disrupted in mutants 2 and 3 are important for this sequence specific activation. Again, quantitative analysis is necessary to make conclusive statements about the exact sequence requirements for activation of gene expression. The preliminary results indicate that the

sequence requirements for activation of gene expression in plant cells by ORCA-3 are the same as for ORCA-1 and 2.

REFERENCES

Allen MD, Yamasaki K, Ohme-Takagi M, Tateno M, Suzuki M (1998) A novel mode of DNA recognition by a β-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. EMBO J 17: 5484-5496.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402.

Bradley JM, Davies KM, Deroles SC et al. The maize *Lc* regulatory gene up-regulates the flavonoid biosynthetic pathway of *Petunia*. Plant J 1998; 13: 381-392.

Gantet P, Imbault, Thiersault M et al. Necessity of a functional octadecanoic pathway for indole alkaloid synthesis by *Catharanthus roseus* cell suspensions cultured in an auxin-starved medium. Plant Cell Physiol 1998; 39: 220-225.

Grotewold E,Chamberlin M, Snook M et al. Engineering secondary metabolism in maize cells by ectopic expression of transcription factors. Plant Cell 1998; 10: 721-740.

Holton TA, Cornish EC. Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 1995; 7:1071-1083.

Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF (1998) *Arabidopsis CBF1* overexpression induces *COR* genes and enhances freezing tolerance. Science 280:104-106.

Lloyd AM, Walbot V, Davis RW. *Arabidopsis* and *Nicotiana* anthocyanin production activated by maize regulators *R* and *C1*. Science 1992; 258:1773-1775.

Martin C, Paz-Ares J (1997) MYB transcription factors in plants. Trends Genet 13:67-73.

Meijer AH, Verpoorte R, Hoge JHC. Regulation of enzymes and genes involved in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. J Plant Res 1993; 3:145-164.

Menke FLH, Parchmann, Mueller MJ, Kijne JW, Memelink J (1999) Involvement of the octadecanoid pathway and protein phosphorylation in fungal-elicitor-induced expression of terpenoid indole alkaloid biosynthetic genes in *Catharanthus roseus*. Plant Physiol: In press.

Mooney M, Desnos T, Harrison K et al. Altered regulation of tomato and tobacco pigmentation genes caused by the *delila* gene of *Antirrhinum*. Plant J 1995; 7: 333-339.

Moreno PRH, van der Heijden R, Verpoorte R. Cell and tissue cultures of *Catharanthus roseus*: A literature survey. Plant Cell Tiss Org Cult 1995; 42:1-25.

Ohme-Takagi M, Shinshi H (1995) Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. Plant Cell 7: 173-182.

Pasquali G, Goddijn OJM, de Waal A, Verpoorte R, Schilperoort RA, Hoge JHC, Memelink J (1994) Coordinate regulation of two indole alkaloid biosynthetic genes from *Catharanthus roseus* by auxin and elicitors. Plant Mol Biol 18: 1121-1131.

Riechmann JL, Meyerowitz EM (1998) The AP2/EREBP family of plant transcription factors. Biological Chemistry 379: 633-646.(review)

Stockinger EJ, Gilmour SJ, Thomashow MF (1997) *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. Proc Natl Acad Sci USA 94: 1035-1040.

Zhou J, Tang X, Martin GB (1997) The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that bind a *cis*-element of pathogenesis-related genes. EMBO J 16: 3207-3218.

CLAIMS

1. A method of modulating in a cell the expression of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor,

the method comprising inserting into the cell a nucleotide sequence coding for a transcription factor other than a transcription factor having a bHLH-type or a MYB-type DNA-binding domain, the nucleotide sequence is operably linked to at least one expression regulating sequence, and/or modifying the expression of a nucleotide sequence coding for such a transcription factor already present in the cell, the transcription factor is capable of regulating the expression of at least one gene coding for a gene product involved in the biosynthesis of said metabolite or precursor, and subjecting the cell to conditions where the inserted nucleotide sequence coding for a transcription factor is expressed or the expression of the nucleotide sequence already present in the cell is modulated.

2. A method according to claim 1 wherein the cell is a plant cell

3. A method according to claim 2 wherein the cell is of a dicotyledonous species.

4. A method according to claim 3 wherein the cell is of a species selected from the group consisting of the *Gentianales* order and the *Cornales* order.

5. A method according to claim 3 wherein the cell is of a species selected from an indole alkaloid producing species.

6. A method according to claim 4 or 5 wherein the cell is a species selected from the group consisting of species belonging to the *Apocynaceae*, *Alangiaceae, Loganiaceae, Icacinaceae, Cornaceae* and *Rubiaceae* families.

7. A method according to claim 6 wherein the cell is of a species belonging to the genus *Catharanthus*.

8. A method according to claim 7 wherein the cell is a *Catharanthus roseus* cell.

9. A method according to claim 2 wherein the plant cell is in the form of a cell culture.

10. A method according to claim 2 wherein the plant cell is in a plant tissue.

11. A method according to claim 1 wherein the nucleotide sequence coding for a transcription factor is derived from a plant cell.

12. A method according to claim 11 wherein the cell is derived from a dicotyledonous species.

13. A method according to claim 12 wherein the cell is derived from a species selected from the group consisting of species of the *Gentianales* order and species of the *Cornales* order.

14. A method according to claim 12 wherein the cell is derived from an indole alkaloid producing species.

15. A method according to claim 13 or 14 wherein the cell is derived from a species selected from the group consisting of species belonging to the *Apocynaceae*, *Alangiaceae*, *Loganiaceae*, *Icacinaceae*, *Cornaceae* and *Rubiaceae* families.

16. A method according to claim 15 wherein the cell is derived from a species belonging to the *Catharanthus* genus.

17. A method according to claim 16 wherein the cell is derived from *Catharanthus roseus*.

18. A method according to claim 1 wherein the nucleotide sequence is a sequence coding for a transcription factor comprising at least one AP2 domain, including at least two AP2 domains and at least three AP2 domains.

19. A method according to claim 18 wherein the nucleotide sequence is a sequence comprising at least one sequence coding for an AP2 domain selected from the group consisting of AP2 domains of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

20. A method according to claim 18 wherein the nucleotide sequence is a sequence coding for an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

21. A method according to claim 18 wherein the nucleotide sequence is a sequence coding for an amino acid sequence comprising an at least one AP2 domain having at least 40% homology, at least 50%, at least 60%, at least 70%, at least 80% or at least 90% homology to any of the at least one AP2 domains of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

22. A method according to claim 18 wherein the nucleotide sequence comprises at least one sequence coding for an AP2 domain of a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4.

23. A method according to claim 18 wherein the nucleotide sequence coding for the at least one AP2 domain is a mutant, allele or variant of a nucleotide sequence coding for an AP2 domain of a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4.

24. A method according to claim 1 wherein the nucleotide sequence coding for a transcription factor is a homologous nucleotide sequence.

25. A method according to claim 1 wherein the nucleotide sequence coding for a transcription factor is a heterologous nucleotide sequence.

26. A method according to claim 24 or 25 wherein the nucleotide sequence is operably linked to a regulating nucleotide sequence with which it is not natively associated.

27. A method according to claim 1 wherein the metabolite is a plant metabolite.

28. A method according to claim 27 wherein the plant metabolite is a secondary metabolite.

29. A method according to claim 28 wherein the secondary metabolite is selected from the group consisting of an alkaloid compound, a phenolic compound and a terpenoid compound.

30. A method according to claim 29 wherein the secondary metabolite is derived from a compound selected from the group consisting of tryptophan, tyrosine, lysine, ornithine, nicotinic acid, anthranilic acid and acetate.

31. A method according to claim 30 wherein the secondary metabolite is an indole alkaloid such as a terpenoid indole alkaloid.

32. A method according to claim 1 wherein the gene product involved in metabolite production is a protein including an enzyme.

33. A method according to claim 32 wherein the enzyme is an enzyme involved in an alkaloid biosynthetic pathway.

34. A method according to claim 33 wherein the enzyme is selected from the group consisting of tryptophan decarboxylase (TDC), geraniol-10-hydroxylase (G10H), strictosidine synthase (STR) and strictosidine glucosidase (SG).

35. A method according to claim 1 wherein the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is enhanced or reduced relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor.

36. A method according to claim 35 wherein the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is enhanced by at least 10%, at least 25% or at least 100%, relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor.

37. A method according to claim 35 wherein the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is reduced by at least 10%, at least 25% or at least 100%, relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor.

38. A method of modulating the stress resistance of a cell, the method comprising inserting into the cell a nucleotide sequence coding for a transcription factor, the nucleotide s quence is operably linked to at least one expression regulating sequence, and/or modifying the expression of a nucleotide sequence coding for such a transcription factor already present in the cell, the transcription factor is capable of regulating the expression of at least one gene coding for a gene product involved in the biosynthesis of said metabolite or precursor, and subjecting the cell to conditions where the inserted nucleotide sequence coding for a transcription factor is expressed or the expression of the nucleotide sequence already present in the cell is modulated, the expression of said nucleotide sequence resulting in a modified responsiveness of the cell towards exogenous stress conditions.

39. A method according to claim 38 wherein the cell is a plant cell

40. A method according to claim 39 wherein the cell is of a dicotyledonous species.

41. A method according to claim 40 wherein the cell is of a species selected from the group consisting of the *Gentianales* order and the *Cornales* order.

42. A method according to claim 40 wherein the cell is of a species selected from an indole alkaloid producing species.

43. A method according to claim 41 or 42 wherein the cell is a species selected from the group consisting of species belonging to the *Apocynaceae*, *Alangiaceae, Loganiaceae, Icacinaceae, Cornaceae* and *Rubiaceae* families.

44. A method according to claim 43 wherein the cell is of a species belonging to the genus *Catharanthus*.

45. A method according to claim 44 wherein the cell is a *Catharanthus roseus* cell.

46. A method according to claim 39 wherein the plant cell is in the form of a cell culture.

47. A method according to claim 39 wherein the plant cell is in a plant tissue.

48. A method according to claim 37 wherein the nucleotide sequence coding for a transcription factor is derived from a plant cell.

49. A method according to claim 48 wherein the cell is derived from a dicotyledonous species.

50. A method according to claim 49 wherein the cell is derived from a species selected from the group consisting of species of the *Gentianales* order and species of the *Cornales* order.

51. A method according to claim 49 wherein the cell is derived from an indole alkaloid producing species.

52. A method according to claim 50 or 51 wherein the cell is derived from a species selected from the group consisting of species belonging to the *Apocynaceae, Alangiaceae, Loganiaceae, Icacinaceae, Cornaceae* and *Rubiaceae* families.

53. A method according to claim 52 wherein the cell is derived from a species belonging to the *Catharanthus* genus.

54. A method according to claim 53 wherein the cell is derived from *Catharanthus roseus*.

55. A method according to claim 38 wherein the nucleotide sequence is a sequence coding for a transcription factor comprising at least one AP2 domain, including at least two AP2 domains and at least three AP2 domains.

56. A method according to claim 38 wherein the nucleotide sequence is a sequence comprising at least one sequence coding for an AP2 domain selected from the group consisting of AP2 domains of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

57. A method according to claim 38 wherein the nucleotide sequence is a sequence coding for an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

58. A method according to claim 38 wherein the nucleotide sequence is a sequence coding for an amino acid sequence comprising at least one AP2 domain having at least 40% homology, including at least 50%, at least 60%, at least 70%, at least 80% or at least 90% homology to any of the at least one AP2 domains of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

59. A method according to claim 38 wherein the nucleotide sequence comprises at least one sequence coding for an AP2 domain of a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4.

60. A method according to claim 38 wherein the nucleotide sequence coding for the at least one AP2 domain is a mutant, allele or variant of a nucleotide sequence coding for an AP2 domain of a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4.

61. A method according to claim 38 wherein the nucleotide sequence coding for a transcription factor is a homologous nucleotide sequence.

62. A method according to claim 38 wherein the nucleotide sequence coding for a transcription factor is a heterologous nucleotide sequence.

63. A method according to claim 61 or 62 wherein the nucleotide sequence is operably linked to a regulating nucleotide sequence with which it is not natively associated.

64. A method according to claim 38 wherein the metabolite is a plant metabolite

65. A method according to claim 64 wherein the plant metabolite is a secondary metabolite.

66. A method according to claim 65 wherein the secondary metabolite is selected from the group consisting of an alkaloid compound, a phenolic compound and a terpenoid compound.

67. A method according to claim 66 wherein the secondary metabolite is derived from a compound selected from the group consisting of tryptophan, tyrosine, lysine, ornithine, nicotinic acid, anthranilic acid and acetate.

67. A method according to claim 66 wherein the secondary metabolite is an indole alkaloid such as a terpenoid indole alkaloid.

69. A method according to claim 38 wherein the gene product involved in metabolite production is a protein including an enzyme.

70. A method according to claim 69 wherein the enzyme is an enzyme involved in an alkaloid biosynthetic pathway.

71. A method according to claim 70 wherein the enzyme is selected from the group consisting of tryptophan decarboxylase (TDC), geraniol-10-hydroxylase (G10H), strictosidine synthase (STR) and strictosidine glucosidase (SG).

72. A method according to claim 38 wherein the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is enhanced or reduced relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor.

73. A method according to claim 72 wherein the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is enhanced by at least 10%, at least 25% or at least 100%, relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor.

74. A method according to claim 72 wherein the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is reduced by at least 10%, at least 25% or at least 100%, relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor.

75. A recombinant cell having, relative to its parent cell, enhanced or reduced biosynthesis of a metabolite or a precursor therefor, and/or enhanced or reduced expression of a gene product involved in metabolite production, the cell comprising a nucleotide sequence

coding for a transcription factor other than a transcription factor having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating the expression of at least one gene coding for a gene product involved in the metabolite production, said sequence is inserted into the cell and/or its expression is modified by operably linking it to a regulating sequence with which it is not natively associated.

76. A cell according to claim 75 which is a plant cell

77. A cell according to claim 76 which is of a dicotyledonous species.

78. A cell according to claim 77 wherein the cell is a cell of a species selected from the group consisting of the *Gentianales* order and the *Cornales* order.

79. A cell according to claim 77 which is a cell of an indole alkaloid producing species.

80. A cell according to claim 78 or 79 which is a cell of a species selected from the group consisting of species belonging to the *Apocynaceae, Alangiaceae, Loganiaceae, Icacinaceae, Cornaceae* and *Rubiaceae* families.

81. A cell according to claim 80 which is a cell of a species belonging to the genus *Catharanthus*.

82. A cell according to claim 81 which is a *Catharanthus roseus* cell.

83. A cell according to claim 76 which is in the form of a cell culture.

84. A cell according to claim 76 wherein the plant cell is in a plant tissue.

85. A cell according to claim 75 wherein the nucleotide sequence coding for a transcription factor is derived from a plant cell.

86. A cell according to claim 85 which is derived from a dicotyledonous species.

87. A cell according to claim 86 which is derived from a species selected from the group consisting of species of the *Gentianales* order and species of the *Cornales* order.

88. A cell according to claim 86 which is derived from an indole alkaloid producing species.

89. A cell according to claim 87 or 88 which is derived from a species selected from the group consisting of species belonging to the *Apocynaceae*, *Alangiaceae, Loganiaceae, Icacinaceae, Cornaceae* and *Rubiaceae* families.

90. A cell according to claim 89 which is derived from a species belonging to the *Catharanthus* genus.

91. A cell according to claim 90 which is derived from *Catharanthus roseus*.

92. A cell according to claim 75 wherein the nucleotide sequence is a sequence coding for a transcription factor comprising at least one AP2 domain, including at least two AP2 domains and at least three AP2 domains.

93. A cell according to claim 85 wherein the nucleotide sequence is a sequence comprising at least one sequence coding for an AP2 domain selected from the group consisting of AP2 domains of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

94. A cell according to claim 93 wherein the nucleotide sequence is a sequence coding for an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

95. A cell according to claim 93 wherein the nucleotide sequence is a sequence coding for an amino acid sequence comprising at least one AP2 domain having at least 40% homology, including at least 50%, at least 60%, at least 70%, at least 80% or at least 90% homology to any of the at least one AP2 domains of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

96. A cell according to claim 75 wherein the nucleotide sequence comprises at least one sequence coding for an AP2 domain of a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4.

97. A cell according to claim 75 wherein the nucleotide sequence coding for the at least one AP2 domain is a mutant, allele or variant of a nucleotide sequence coding for an AP2 domain of a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4.

98. A cell according to claim 75 wherein the nucleotide sequence coding for a transcription factor is a homologous nucleotide sequence.

99. A cell according to claim 75 wherein the nucleotide sequence coding for a transcription factor is a heterologous nucleotide sequence.

100. A cell according to claim 98 or 99 wherein the nucleotide sequence is operably linked to a regulating nucleotide sequence with which it is not natively associated.

101. A cell according to claim 75 wherein the metabolite is a plant metabolite

102. A cell according to claim 101 wherein the plant metabolite is a secondary metabolite.

103. A cell according to claim 102 wherein the secondary metabolite is selected from the group consisting of an alkaloid compound, a phenolic compound and a terpenoid compound.

104. A method according to claim 103 wherein the secondary metabolite is derived from a compound selected from the group consisting of tryptophan, tyrosine, lysine, ornithine, nicotinic acid, anthranilic acid and acetate.

105. A cell according to claim 103 wherein the secondary metabolite is an indole alkaloid such as a terpenoid indole alkaloid.

106. A cell according to claim 75 wherein the gene product involved in metabolite production is a protein including an enzyme.

107. A cell according to claim 106 wherein the enzyme is an enzyme involved in an alkaloid biosynthetic pathway.

108. A cell according to claim 107 wherein the enzyme is selected from the group consisting of tryptophan decarboxylase (TDC), geraniol-10-hydroxylas (G10H), strictosidine synthase (STR) and strictosidine glucosidas (SG).

109. A cell according to claim 75 where the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is enhanced or reduced relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor.

110. A cell according to claim 109 where the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is enhanced by at least 10%, at least 25% or at least 100%, relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor.

111. A cell according to claim 110 wherein at least one pathway leading to the biosynthesis of at least one precursor for the metabolite can be stimulated in said cell in the presence of an inducing agent.

112. A cell according to claim 109 wherein the expression in the cell of one or more genes involved in the biosynthesis of a metabolite or a precursor herefor is reduced by at least 10%, at least 25% or at least 100%, relative to a cell into which a nucleotide sequence coding for a transcription factor is not inserted or which is without modified expression of the already present nucleotide sequence coding for a transcription factor.

113. A method of producing a metabolite including a plant secondary metabolite, the method comprising providing a recombinant cell according to any of claims 75-112, cultivating said cell under conditions where the nucleotide sequence coding for the transcription factor regulating the expression of at least one gene coding for a gene product involved in the metabolite production is expressed, and recovering the metabolite.

114. A method according to claim 113 wherein the cell is cultivated under conditions where at least one precursor for the metabolite is added in a form which can be assimilated by said cell.

115. A method of constructing a recombinant cell according to any of claims 75-112, said method comprising the steps of (i) identifying in a source cell a nucleotide sequence coding for a transcription factor other than a transcription factor having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating in the source cell the expression of at least one gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, (ii) isolating said nucleotide sequence and (iii) inserting said isolated nucleotide sequence into a host cell comprising a gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, the insertion of said nucleotide sequence resulting in that the expression of the gene of the host cell is modulated relative to the expression level in a host cell not comprising the inserted nucleotide sequence.

116. A method according to claim 115 wherein at least two isolated nucleotide sequences are inserted into the host cell.

117. An isolated nucleic acid molecule comprising a nucleotide sequence coding for an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

118. An isolated nucleic acid according to claim 117 comprising a nucleotide sequence which is selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4 and a mutant, allele or variant hereof.

119. A polypeptide which includes an amino acid sequence having at least 40% homology including at least 50%, such as at least 60%, at least 70%, at least 80% and at least 90% to an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

120. A polypeptide having the amino acid sequence shown in SEQ ID NO:5, SEQ ID NO:6 or SEQ ID NO:7.

121. A vector comprising a nucleic acid molecule according to claim 117 or 118.

122. A vector according to claim 117 comprising at least one further sequence coding for a transcription factor.

123. A host cell comprising a vector according to claim 117 or 118.

124. A cell according to claim 123 which is selected from the group consisting of a bacterium, a fungal cell, a yeast cell, an animal cell, a unicellular eukaryotic cell, an algae cell and a plant cell.

125. A method of isolating a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating in a cell the expression of at least one gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, the method comprising the steps of

(i) probing a cDNA or genomic library with a probe comprising at least a fragment of a nucleic acid according to claim 117 or 118;

(ii) identifying a DNA clone that hybridises with said at least a fragment of a nucleic acid; and

(iii) isolating the DNA clone identified in step (ii) wherein the nucleic acid sequence is coding for all or a part of said transcription factor.

126. A method of isolating a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating in a cell the expression of at least one gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, the method comprising the steps of

(i) synthesising an oligonucleotide primer set corresponding to at least a fragment of a sequence selected from the group consisting of SEQ ID NO:1, 2, 3 and 4; and

(ii) amplifying cDNA or genomic DNA using said primer set in a polymerase chain reaction wherein the amplified nucleic acid fragment is coding for all or a part of said transcription factor

to obtain a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain.

127. An isolated nucleic acid which is a nucleotide sequence as shown in SEQ ID NO: 8 and a mutant, allele or variant hereof.

128. Use of an isolated nucleic acid sequence according to claim 127 for isolating a nucleotide sequence coding for a transcription factor not having a bHLH-type or a MYB-type DNA-binding domain, said transcription factor is capable of regulating in a cell the expression of at least one gene coding for a gene product involved in biosynthesis of a metabolite or a precursor therefor, the method comprising the steps as essentially described in Example 2

ABSTRACT

A method for modulating in a cell the expression of one or more genes involved in the biosynthesis of a metabolite or a precursor for this metabolite by inserting into a cell a nucleotide sequence coding for a transcription factor comprising an AP2 DNA-binding domain, and/or by modifying the expression of a nucleotide sequence coding for such a transcription factor already present in the cell is provided. The method is useful for enhancing the biosynthesis of secondary metabolites in plants such as alkaloids including terpenoid indole alkaloids. Also provided is a method for enhancing stress tolerance in plants by the use of such transcription factors.

SEQUENCE LISTING

<110> Johan Memelink
Leslie van der Fits
Frank L.H. Menke
Jan W. Kijne

<120> METHOD OF MODULATING METABOLITE BIOSYNTHESIS IN RECOMBINANT CELLS

<130> 22209xx1

<160> 8

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2298
<212> DNA
<213> Catharanthus roseus

<400> 1

atcaaatttc acacagtttt agaactctac gacctatttg ttactgaaaa ttactggaat 60
tactaaaatc ggaagaagaa atcaacgcga cgaaagagaa aaagaacaaa agggtttcgt 120
ttttgtaaag tttgattctt ggcggagatt ttcgacaaag gagtgggcaa tttgtgcaat 180
acttctgaga aaattgaaag agatacaagg atggctcttc ttgatcaggc atccaatttg 240
agtcccatgc cttttgattt cactagaaag aggaagtcga ggaggaggga tggtactaag 300
aacgtagcgg agacacttgc aaagtggaaa gagtataatg agaaacttga tgctttagat 360
ggagggaaac cagctcggaa ggttcctgcc aaaggatcaa aaaagggatg tatgaaaggt 420
aaaggaggcc ctgagaattc tcactgcaaa tacagaggag ttaggcagag gacatggggt 480
aaatgggtgg ccgaaattcg ggaaccaaac aggggtagca ggctttggtt gggtacattc 540
agaaacgcga tagaagctgc acttgcttat gatgaagcag cgagggcgat gtatgggcct 600
tgtgctaggc ttaatcttcc gaactatagg gcttcagaag aatcttcttc cttgccaaca 660
acatcaggat cagatacgac tactgcttct ggcatctcag aggtctctgt ctatgaagac 720
aaaaagttca caccagttgt ttccggattg aaacaagatg acaagggtga atcattagag 780
tcagctgata gtaaacctca actcctggtc gatgctggca ctcccatgag tgcagtgaag 840
gaagaaccaa aagaatatca ggttatggat tcccagtctg aagggcaatt cggagacgag 900
gaaccgccta gcaagcttgt ttgtaaagaa gtcgactttg ggcaggatca agctgttgtt 960
cctgctgtta aaaatgctga ggagatgggt ggagagatgg gtggagatat actgaaaggc 1020
tgttctttgt ctgagatgtt tgatgtggac gagttgttga gcgttttaga ttctacaccc 1080
ctccatgcct cagattttca gcatggcatg ggaaatggta atgtaaaggc agaggctgct 1140
tacaactatg ctccttcatg ggactcggcc ttccagttgc agaatcaaga tcctaagcta 1200
ggaagtcagc agcacatggc gcagacaccc ccagaaatta attccgggct tgatttttg 1260
cagccaggaa gacaagagga ctcctatttt actttgggtg atctagactt tcttgatttg 1320
ggtgctgaat tgggattgta aatccgaagt tgttgaagct aaaagcggcg actatgaaac 1380
tggaattttg gaacggctta ttgttcctgg tgtttgtctt agttctagtc tgtttatgta 1440
ctagaacttg acatatagga ggcttttgaa agctgaacaa acgaagtgtg aattattttc 1500
ttttttgtt tttctgcagc gatgtatact aacatctcta ctactaaaat tacgtctctt 1560
cgtcttcact aacagtaggg tggagctgat tctcttttaa gtttttcaga aggggaattc 1620
agctatgagt ttagaggcag ggtagtgtag ttcagtgagc agattctttc tgtagatatc 1680
tctagtcttt tggtttcttg gaatgttttt tctggtggaa taaaggaagc tgtgttcttg 1740
tgggatttgg ttcttattag aatcacccct aatttgatca aacagctcca gaacgctcat 1800
tctcttccca atcagcctat gcatgaaatt cttctcttct ctctcccctc tcatggccga 1860
aaattgcaag agaaaaagag ccatctttag ctctttagtt tcttttcaat ttcttctctt 1920
cttttctttg tttttcttcg tcggaagtac tcatgcgggt gagagaatac tccagaagat 1980
catggagttt tgtggaagta ttaagcaatt tcaatggagg tctcaggaac aaatcgcgac 2040
ttctccggtc tgcgcgtttc ttcctctccg gcaatgtcag gtgactcttg ttgggctcac 2100
ggtggttagc gtcttatggg gaagctcacg atgaaaatgc atggcgagct gaagattttt 2160
ttttttttt ttttttttga tgaaaatgct tggccagcaa aattctggcg agccaatttt 2220
tgcttttctt tctttatatg aaatccttgg accatcaagc tgttcttttt tttaacaaca 2280
taatgtgtta gccagcct 2298

<210> 2
<211> 1754
<212> DNA
<213> Catharanthus roseus

<400> 2
tgtaaatcaa atttcacaca gttttagaac tctacgacct atttgttact gaaaattact 60
ggaattacta aaatcggaag aagaaatcaa cgcgacgaaa gagaaaaaga acaaaagggt 120
ttcgtttttg taaagtttga ttcttggcgg agattttcga caaaggagtg ggcaatttgt 180
gcaatacttc tgagaaaatt gaaagagata caaggatggc tcttcttgat caggcatcca 240
atttgagtcc catgcctttt gatttcacta gaaagaggaa gtcgaggagg agggatggta 300
ctaagaacgt agcggagaca cttgcaaagt ggaaagagta taatgagaaa cttgatgctt 360
tagatggagg gaaaccagct cggaaggttc ctgccaaagg atcaaaaaag ggatgtatga 420
aaggtaaagg aggccctgag aattctcact gcaaatacag aggagttagg cagaggacat 480
ggggtaaatg ggtggccgaa attcgggaac caaacagggg tagcaggctt tggttgggta 540
cattcagaaa cgcgatagaa gctgcacttg cttatgatga agcagcgagg gcgatgtatg 600
ggccttgtgc taggcttaat cttccgaact atagggcttc agaagaatct tcttccttgc 660
caacaacatc aggatcagat acgactactg cttctggcat ctcagaggtc tctgtctatg 720
aagacaaaaa gttcacacca gttgtttccg gattgaaaca agatgacaag ggtgaatcat 780
tagagtcagc tgatagtaaa cctcaactcc tggtcgatgc tggcactccc atgagtgcag 840
tgaaggaaga accaaaagaa tatcaggtta tggattccca gtctgaaggg caattcggag 900
acgaggaacc gcctagcaag cttgtttgta aagaagtcga ctttgggcag gatcaagctg 960
ttgttcctgc tgttaaaaat gctgaggaga tgggtggaga gatgggtgga gatatactga 1020
aaggctgttc tttgtctgag atgtttgatg tggacgagtt gttgagcgtt ttagattcta 1080
caccctcca tgcctcagat tttcagcatg gcatgggaaa tggtaatgta aaggcagagg 1140
ctgcttacaa ctatgctcct tcatgggact cggccttcca gttgcagaat caagatccta 1200
agctaggaag tcagcagcac atggcgcaga caccccaga aattaattcc gggcttgatt 1260
ttttgcagcc aggaagacaa gaggactcct attttacttt gggtgatcta gactttcttg 1320
atttgggtgc tgaattggga ttgtaaatcc gaagttgttg aagctaaaag cggcgactat 1380
gaaactggaa ttttggaacg gcttattgtt cctggtgttt gtcttagttc tagtctgttt 1440
atgtactaga acttgacata taggaggctt ttgaaagctg aacaaacgaa gtgtgaatta 1500
ttttcttttt ttgtttttct gcagcgatgt atactaacat ctctactact aaaattacgt 1560
ctcttcgtct tcactaacag tagggtggag ctgattctct tttaagtttt tcagaagggg 1620
aattcagcta tgagtttaga ggcagggtag tgtagttcag tgagcagatt ctttctgtag 1680
atatctctag tcttttggtt tcttggaatg ttttttctgg tggaataaag atggcatagg 1740
tggaggttgt atct 1754

<210> 3
<211> 885
<212> DNA
<213> Catharanthus roseus

<400> 3
caacaataat gtatcaatca aatgcccata attccgatca tctaaccttc ttaccacctt 60
tagtagatta tcaattcctc aacaacgatt ttgatttttc agaaatattt acagatttca 120
attacgctaa ttataattat aatacttcta cctcagataa tttctctggt tttcaattca 180
atgaaaattg cgaagaaatt atttcaccaa attatgcttc ggaagattta tcggatatta 240
ttttaacaga tattttcaag gatcaggata attacgaaga cgaagtcgtt gcgggagaac 300
aagaagaaga attaattacg acacctacct ctcgcggcgg cggcggcggc ggatgtgagc 360
agagatcgaa tgaggaatgg attaggtacc gtggcgttag acggcggcca tgggggaaat 420
tcgctgcgga aatcagggat cccaagagaa aaggatcgag gatatggttg ggaacttacg 480
agacggcgga agatgcggca ttagctttcg atcaagcggc gtttcaactc cgtggttcta 540
gagctaggtt aaattttccc aatcttattg gttctgctaa tgctccggtt agagtaagtc 600
ctagacgccg atcttcatcg tgtcatcttc gtcctcaata atcctatcca cagttccatg 660
gggatagtaa atttttttctt tgagtttttt agaagttata ttatctattg aaaaaataca 720
aaacattgca aatatttttt tagtacgtct ctatacttct ttttagtaat attcggatca 780
tgagcatggg gaaggtgata ttatccattg tcataaatta atagatacag tatcataaat 840
taatatgtac gaattacaag taaaatatag taagtgttaa tattg 885

<210> 4
<211> 792
<212> DNA
<213> Catharanthus roseus

<400> 4

```
ttctaaaaaa gaagaaaaat gtccgaagaa atcatttccg tctcagatcg atttcttctt      60
tccttaatcg aagaacatct tctcagcgat aattctgatg attccagctc ggaattgact     120
tctacagagg aaaattggga agaaattttt gcagattttc taaattggtc gggatccgaa     180
atacagaaac gcggtagccc gagttccgaa agctgtcaat cgaattcaat ggcggaaagc     240
tgtcaggagg attctgttgt gggaaccccg ccagaagcgg cggccggagg aggttgttcg     300
aaggattgga accggtataa gggcgttaga cggcggccgt gggggaagtt cgcggcggag     360
ataagggatc cgaaaaagaa aggatccagg atttggttgg gtacatacga gacacctgag     420
gatgcagcat tggcttatga tgcagccgcg tttaatatgc gtggagctaa agctaggctt     480
aattttcctc atttgattgg ttcgaatatt tccggacccg ttagagtaaa cccgagaaaa     540
cgtttccctg cggagccttc tacgacgtcg tcgtcttctt cttcttcttc gtctgaaaat     600
agtggaggaa ggaagaagag acgatattaa ttaattatta aaagtggagg attaaaaaaa     660
ttctgtgaaa tgagagatta ttacgtggtt tttgttaagc ccgataatcc ctcattgtaa     720
aattattaac ttcatcgatg ttctttttta aatctttgga atgtacaaaa ttttatatcc     780
aaaaaagttc ac                                                        792
```

<210> 5
<211> 376
<212> PRT
<213> Catharanthus roseus

<400> 5

```
Met Ala Leu Leu Asp Gln Ala Ser Asn Leu Ser Pro Met Pro Phe Asp
 1               5                  10                  15
Phe Thr Arg Lys Arg Lys Ser Arg Arg Arg Asp Gly Thr Lys Asn Val
            20                  25                  30
Ala Glu Thr Leu Ala Lys Trp Lys Glu Tyr Asn Glu Lys Leu Asp Ala
        35                  40                  45
Leu Asp Gly Gly Lys Pro Ala Arg Lys Val Pro Ala Lys Gly Ser Lys
    50                  55                  60
Lys Gly Cys Met Lys Gly Lys Gly Gly Pro Glu Asn Ser His Cys Lys
65                  70                  75                  80
Tyr Arg Gly Val Arg Gln Arg Thr Trp Gly Lys Trp Val Ala Glu Ile
                85                  90                  95
Arg Glu Pro Asn Arg Gly Ser Arg Leu Trp Leu Gly Thr Phe Arg Asn
            100                 105                 110
Ala Ile Glu Ala Ala Leu Ala Tyr Asp Glu Ala Ala Arg Ala Met Tyr
        115                 120                 125
Gly Pro Cys Ala Arg Leu Asn Leu Pro Asn Tyr Arg Ala Ser Glu Glu
    130                 135                 140
Ser Ser Ser Leu Pro Thr Thr Ser Gly Ser Asp Thr Thr Thr Ala Ser
145                 150                 155                 160
Gly Ile Ser Glu Val Ser Val Tyr Glu Asp Lys Lys Phe Thr Pro Val
                165                 170                 175
Val Ser Gly Leu Lys Gln Asp Asp Lys Gly Glu Ser Leu Glu Ser Ala
            180                 185                 190
Asp Ser Lys Pro Gln Leu Leu Val Asp Ala Gly Thr Pro Met Ser Ala
        195                 200                 205
Val Lys Glu Glu Pro Lys Glu Tyr Gln Val Met Asp Ser Gln Ser Glu
    210                 215                 220
Gly Gln Phe Gly Asp Glu Glu Pro Pro Ser Lys Leu Val Cys Lys Glu
225                 230                 235                 240
Val Asp Phe Gly Gln Asp Gln Ala Val Val Pro Ala Val Lys Asn Ala
                245                 250                 255
Glu Glu Met Gly Gly Glu Met Gly Gly Asp Ile Leu Lys Gly Cys Ser
            260                 265                 270
Leu Ser Glu Met Phe Asp Val Asp Glu Leu Leu Ser Val Leu Asp Ser
        275                 280                 285
Thr Pro Leu His Ala Ser Asp Phe Gln His Gly Met Gly Asn Gly Asn
    290                 295                 300
Val Lys Ala Glu Ala Ala Tyr Asn Tyr Ala Pro Ser Trp Asp Ser Ala
305                 310                 315                 320
Phe Gln Leu Gln Asn Gln Asp Pro Lys Leu Gly Ser Gln Gln His Met
                325                 330                 335
```

```
Ala Gln Thr Pro Pro Glu Ile Asn Ser Gly Leu Asp Phe Leu Gln Pro
            340                 345                 350
Gly Arg Gln Glu Asp Ser Tyr Phe Thr Leu Gly Asp Leu Asp Phe Leu
        355                 360                 365
Asp Leu Gly Ala Glu Leu Gly Leu
    370                 375
```

<210> 6
<211> 210
<212> PRT
<213> Catharanthus roseus

<400> 6

```
Met Tyr Gln Ser Asn Ala His Asn Ser Asp His Leu Thr Phe Leu Pro
 1               5                  10                  15
Pro Leu Val Asp Tyr Gln Phe Leu Asn Asn Asp Phe Asp Phe Ser Glu
            20                  25                  30
Ile Phe Thr Asp Phe Asn Tyr Ala Asn Tyr Asn Tyr Asn Thr Ser Thr
        35                  40                  45
Ser Asp Asn Phe Ser Gly Phe Gln Phe Asn Glu Asn Cys Glu Glu Ile
    50                  55                  60
Ile Ser Pro Asn Tyr Ala Ser Glu Asp Leu Ser Asp Ile Ile Leu Thr
65                  70                  75                  80
Asp Ile Phe Lys Asp Gln Asp Asn Tyr Glu Asp Glu Val Val Ala Gly
                85                  90                  95
Glu Gln Glu Glu Glu Leu Ile Thr Thr Pro Thr Ser Arg Gly Gly Gly
            100                 105                 110
Gly Gly Gly Cys Glu Gln Arg Ser Asn Glu Glu Trp Ile Arg Tyr Arg
        115                 120                 125
Gly Val Arg Arg Arg Pro Trp Gly Lys Phe Ala Ala Glu Ile Arg Asp
    130                 135                 140
Pro Lys Arg Lys Gly Ser Arg Ile Trp Leu Gly Thr Tyr Glu Thr Ala
145                 150                 155                 160
Glu Asp Ala Ala Leu Ala Phe Asp Gln Ala Ala Phe Gln Leu Arg Gly
                165                 170                 175
Ser Arg Ala Arg Leu Asn Phe Pro Asn Leu Ile Gly Ser Ala Asn Ala
            180                 185                 190
Pro Val Arg Val Ser Pro Arg Arg Arg Ser Ser Ser Cys His Leu Arg
        195                 200                 205
Pro Gln
    210
```

<210> 7
<211> 203
<212> PRT
<213> Catharanthus roseus

<400> 7

```
Met Ser Glu Glu Ile Ile Ser Val Ser Asp Arg Phe Leu Leu Ser Leu
 1               5                  10                  15
Ile Glu Glu His Leu Leu Ser Asp Asn Ser Asp Asp Ser Ser Ser Glu
            20                  25                  30
Leu Thr Ser Thr Glu Glu Asn Trp Glu Glu Ile Phe Ala Asp Phe Leu
        35                  40                  45
Asn Trp Ser Gly Ser Glu Ile Gln Lys Arg Gly Ser Pro Ser Ser Glu
    50                  55                  60
Ser Cys Gln Ser Asn Ser Met Ala Glu Ser Cys Gln Glu Asp Ser Val
65                  70                  75                  80
Val Gly Thr Pro Pro Glu Ala Ala Ala Gly Gly Gly Cys Ser Lys Asp
                85                  90                  95
Trp Asn Arg Tyr Lys Gly Val Arg Arg Arg Pro Trp Gly Lys Phe Ala
            100                 105                 110
Ala Glu Ile Arg Asp Pro Lys Lys Lys Gly Ser Arg Ile Trp Leu Gly
        115                 120                 125
```

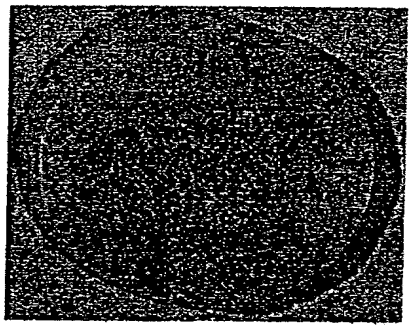
```
Thr Tyr Glu Thr Pro Glu Asp Ala Ala Leu Ala Tyr Asp Ala Ala Ala
    130                 135                 140
Phe Asn Met Arg Gly Ala Lys Ala Arg Leu Asn Phe Pro His Leu Ile
145                 150                 155                 160
Gly Ser Asn Ile Ser Gly Pro Val Arg Val Asn Pro Arg Lys Arg Phe
                165                 170                 175
Pro Ala Glu Pro Ser Thr Thr Ser Ser Ser Ser Ser Ser Ser Ser Ser
            180                 185                 190
Glu Asn Ser Gly Gly Arg Lys Lys Arg Arg Tyr
        195                 200
```

<210> 8
<211> 48
<212> DNA
<213> Catharanthus roseus

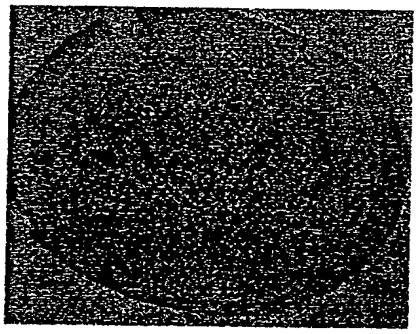
<400> 8
gtacatcact cttagaccgc cttctttgaa agtgatttcc cttggacc 48

4RV - *gusA*
co-bombarded with

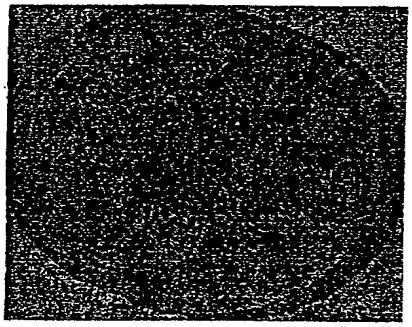
empty vector



35S - *Orca1*



35S - *Orca2*



35S - *Orca3*

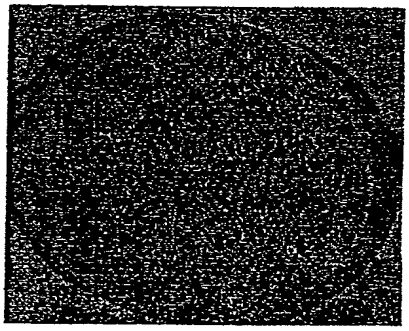


Fig. 12

Tdc - promoter - *gusA*

co-bombarded with

empty vector   35S - *Orca3*

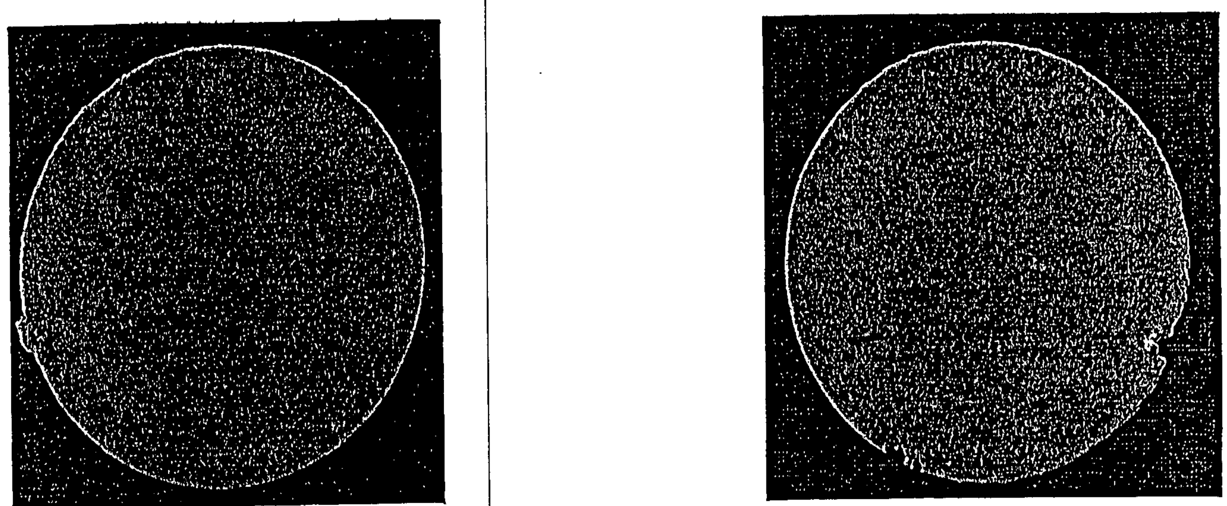
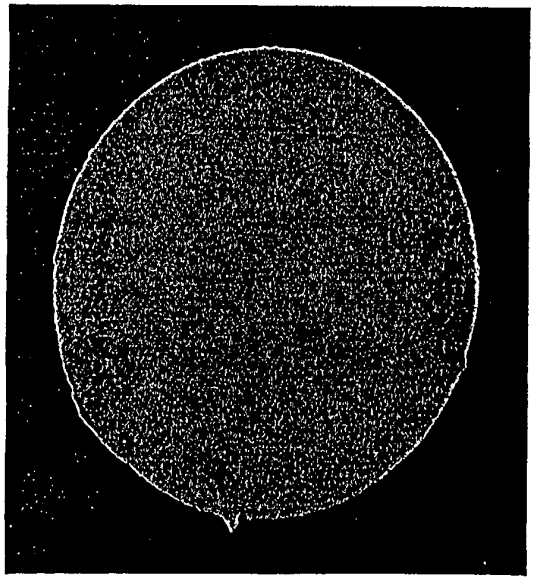


Fig. 11

Str **- promoter -** ***gusA***

co-bombarded with

empty vector



35S - ***Orca3***

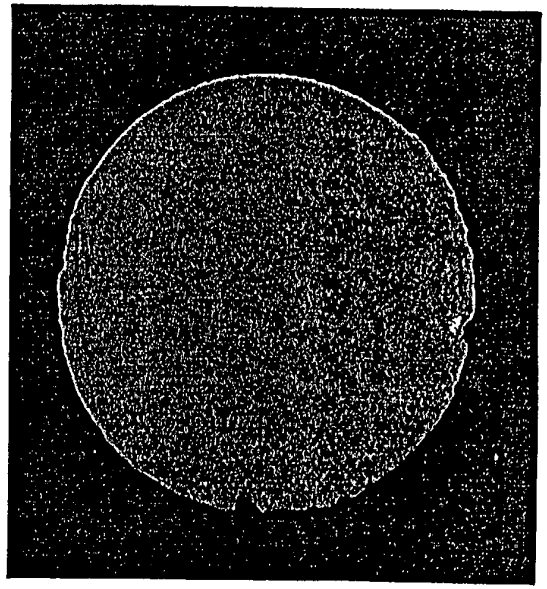


Fig. 10

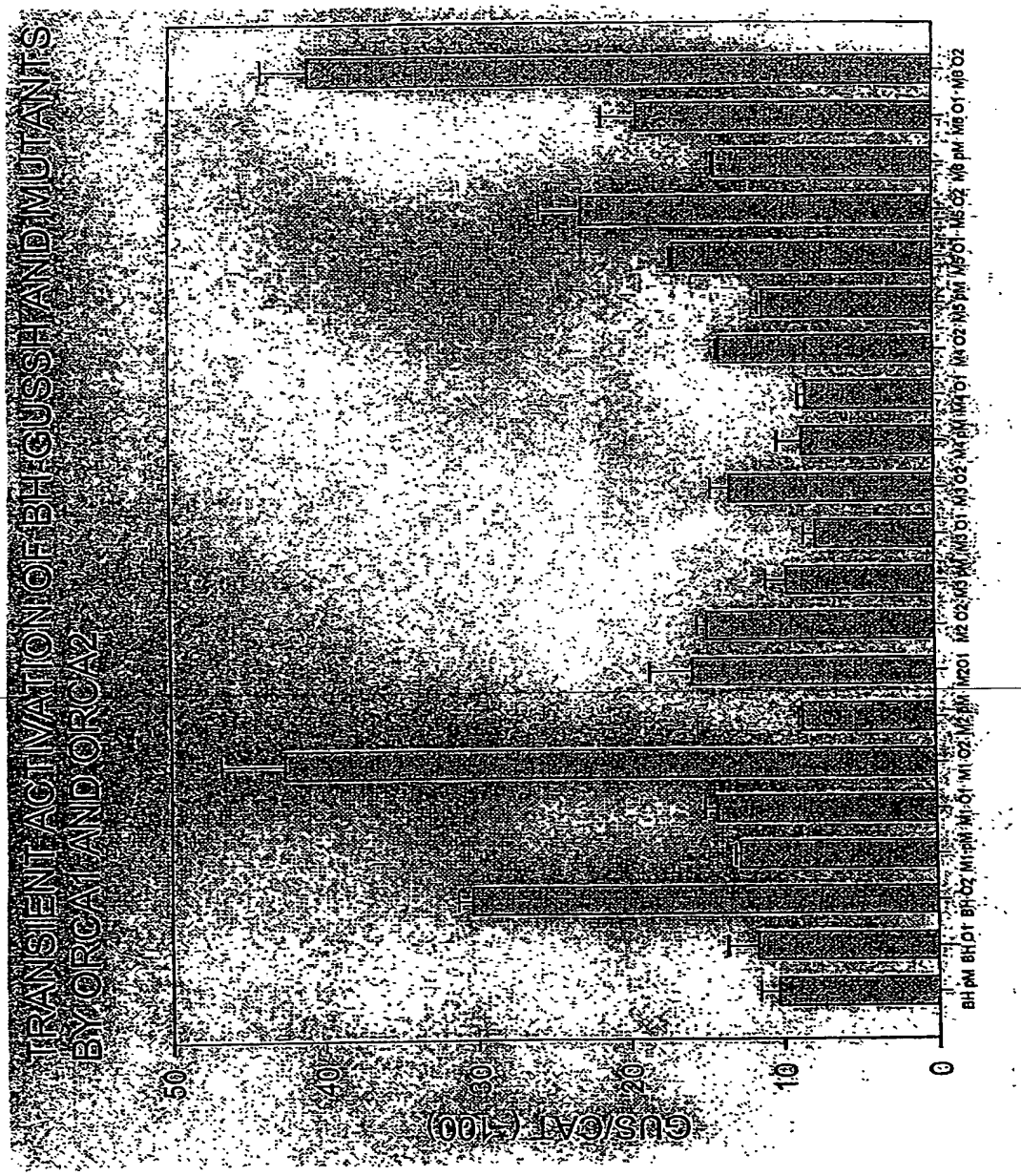

TRANSIENT ACTIVATION OF BHLHB2 USSH AND MUTANTS
BY ORCA1 AND ORCA2
GUS/CAT (*100)
50
40
30
20
10
0
BH pM
BH O1
BH O2
M1 pM
M1 O1
M1 O2
M2 pM
M2 O1
M2 O2
M3 pM
M3 O1
M3 O2
M4 pM
M4 O1
M4 O2
M5 pM
M5 O1
M5 O2
M6 pM
M6 O1
M6 O2

Fig. 9

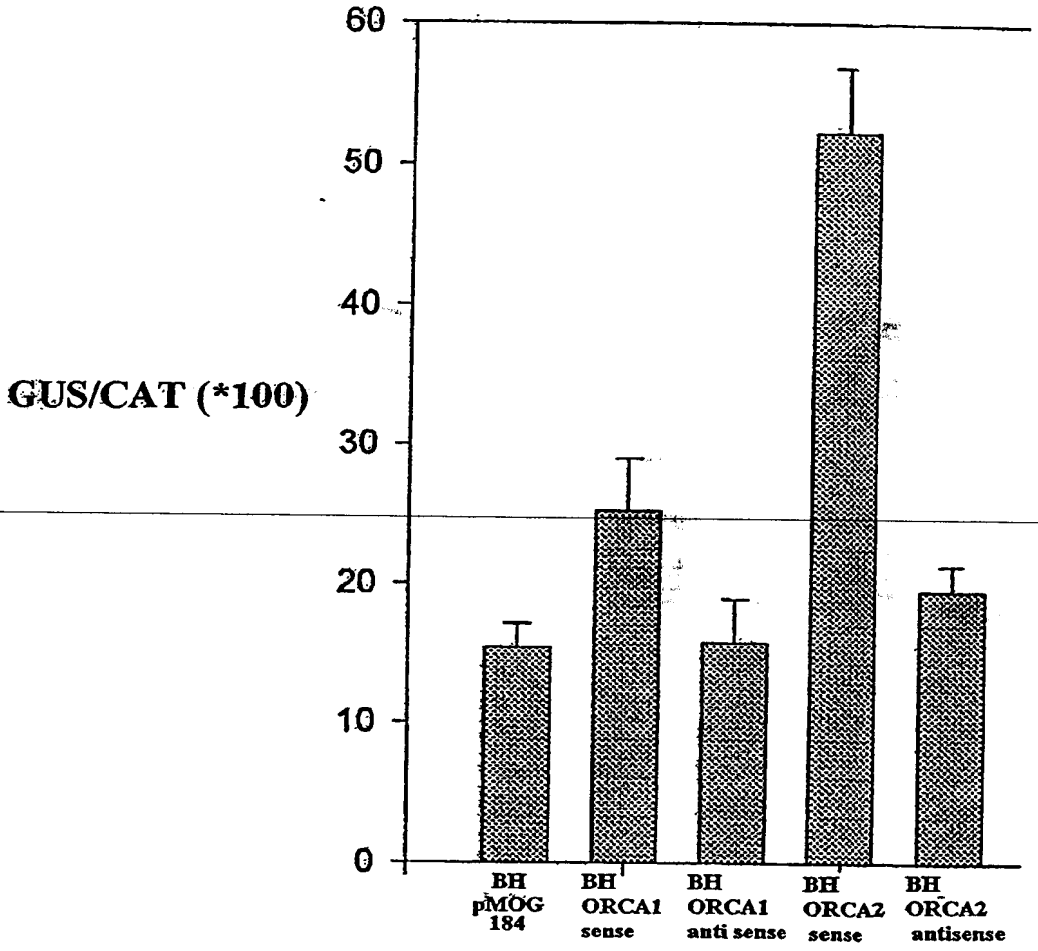

GUS/CAT (*100)
60
50
40
30
20
10
0
BH pMOG 184
BH ORCA1 sense
BH ORCA1 anti sense
BH ORCA2 sense
BH ORCA2 antisense

Fig. 8

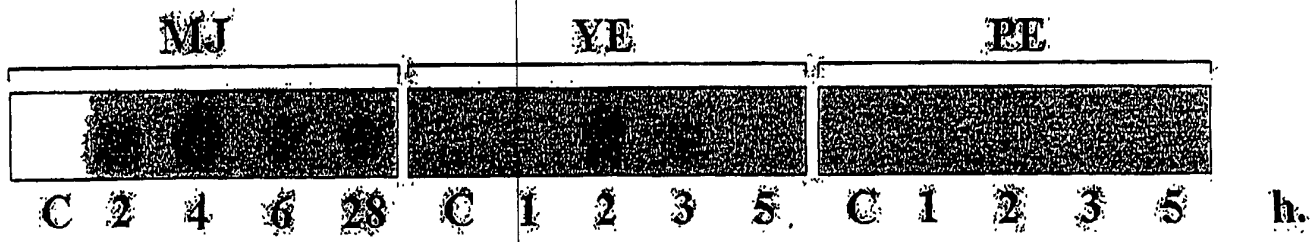


Fig. 7

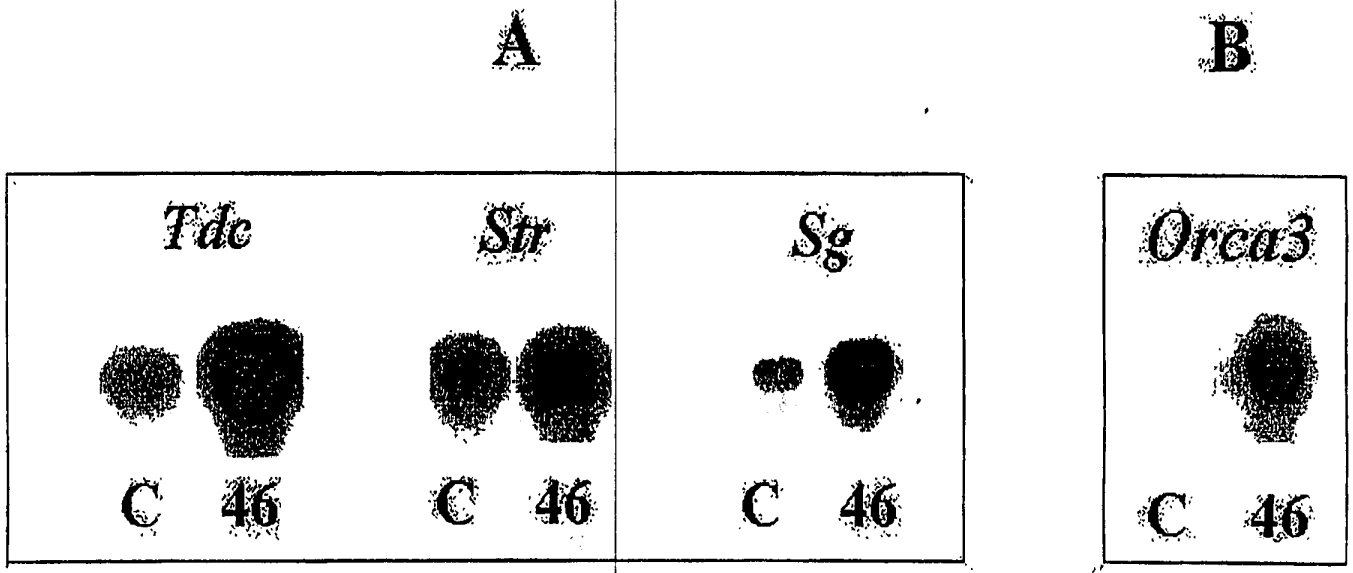


Fig. 6

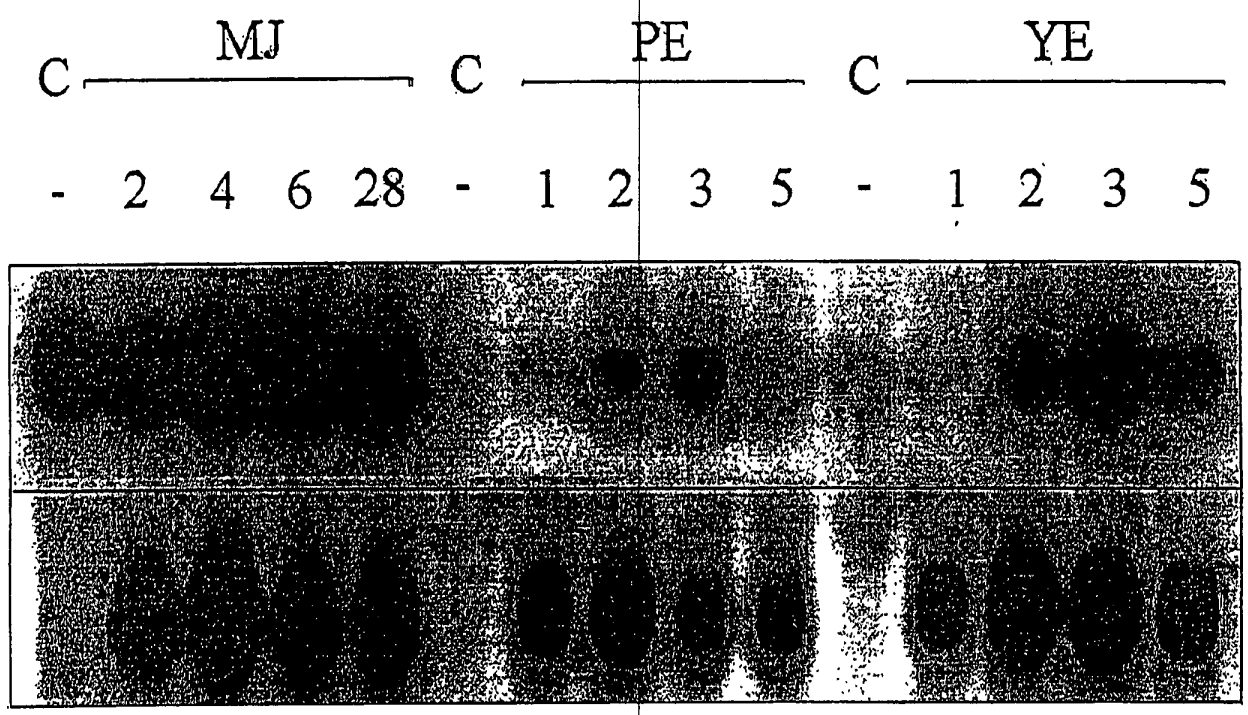


Fig. 5

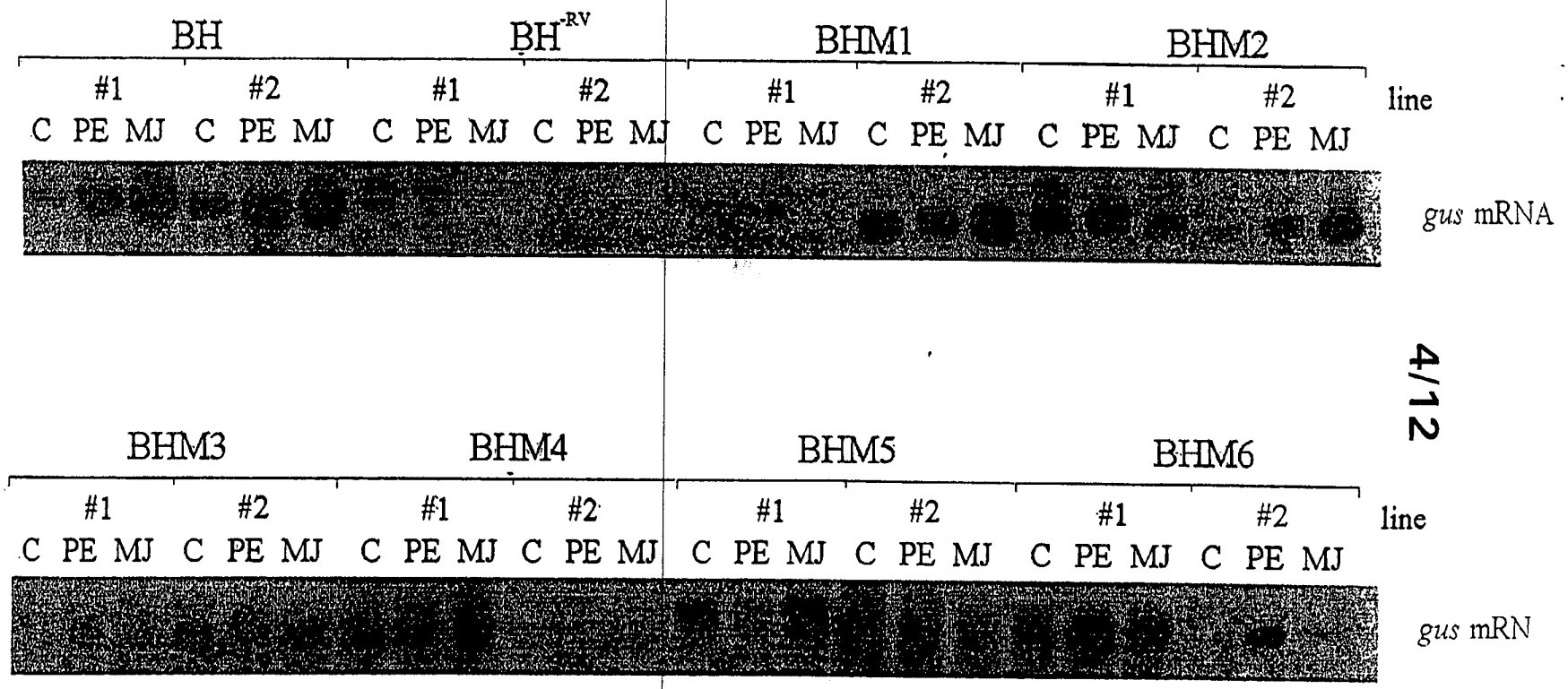

BH
BH-RV
BHM1
BHM2
#1
#2
C PE MJ
line
gus mRNA
BHM3
BHM4
BHM5
BHM6
gus mRN

Fig. 4

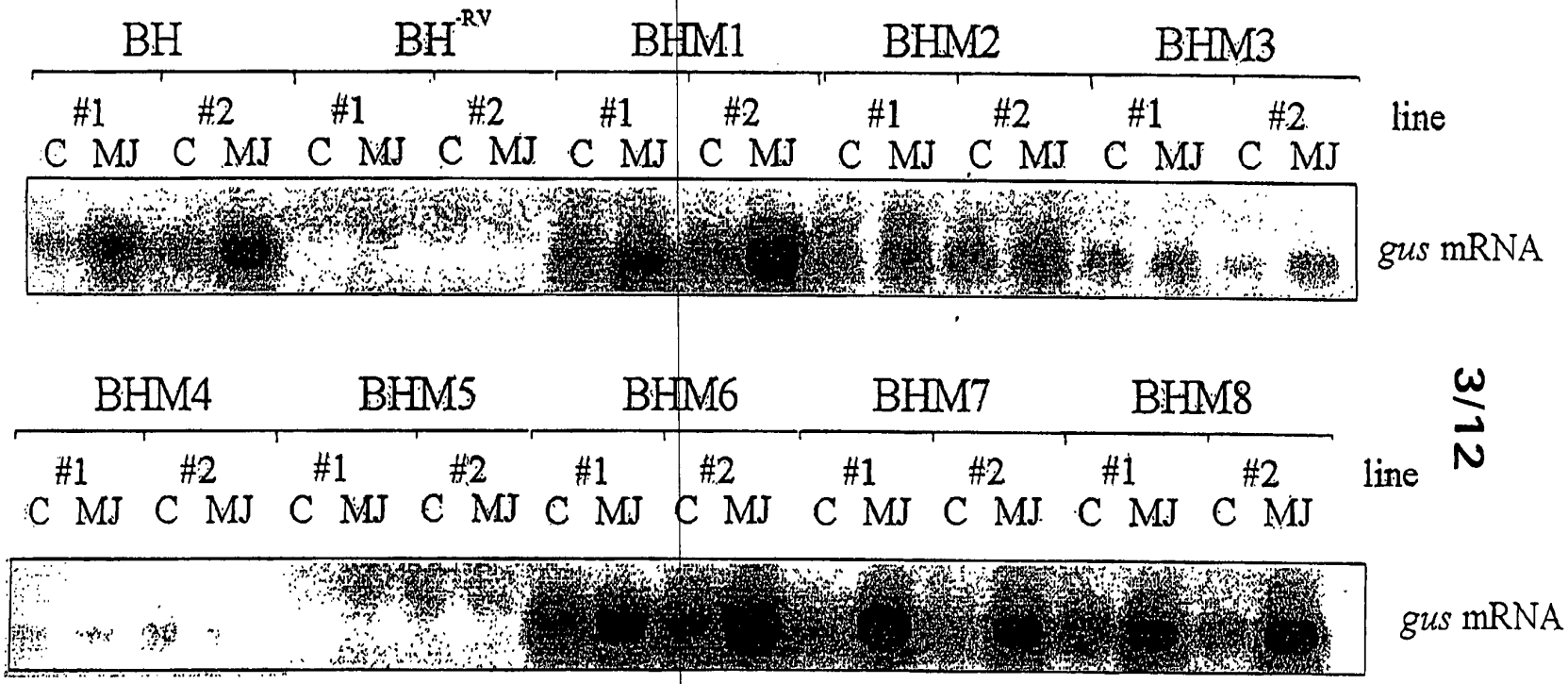

BH
BH-RV
BHM1
BHM2
BHM3
#1 #2 #1 #2 #1 #2 #1 #2 #1 #2
C MJ C MJ C MJ C MJ C MJ C MJ C MJ C MJ C MJ C MJ
line
gus mRNA
BHM4
BHM5
BHM6
BHM7
BHM8
#1 #2 #1 #2 #1 #2 #1 #2 #1 #2
C MJ C MJ C MJ C MJ C MJ C MJ C MJ C MJ C MJ C MJ
line
gus mRNA

Fig. 3

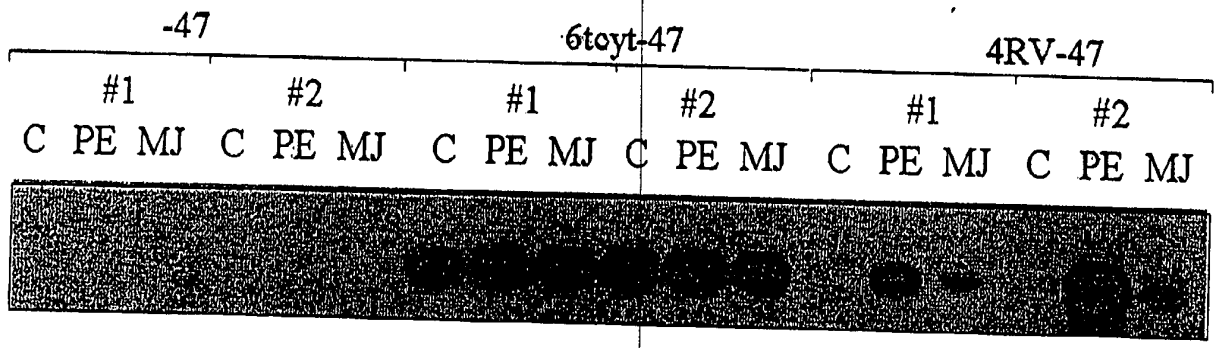

-47
6toyt-47
4RV-47
#1
#2
#1
#2
#1
#2
C PE MJ C PE MJ C PE MJ C PE MJ C PE MJ C PE MJ
line
gus mRNA

Fig. 2

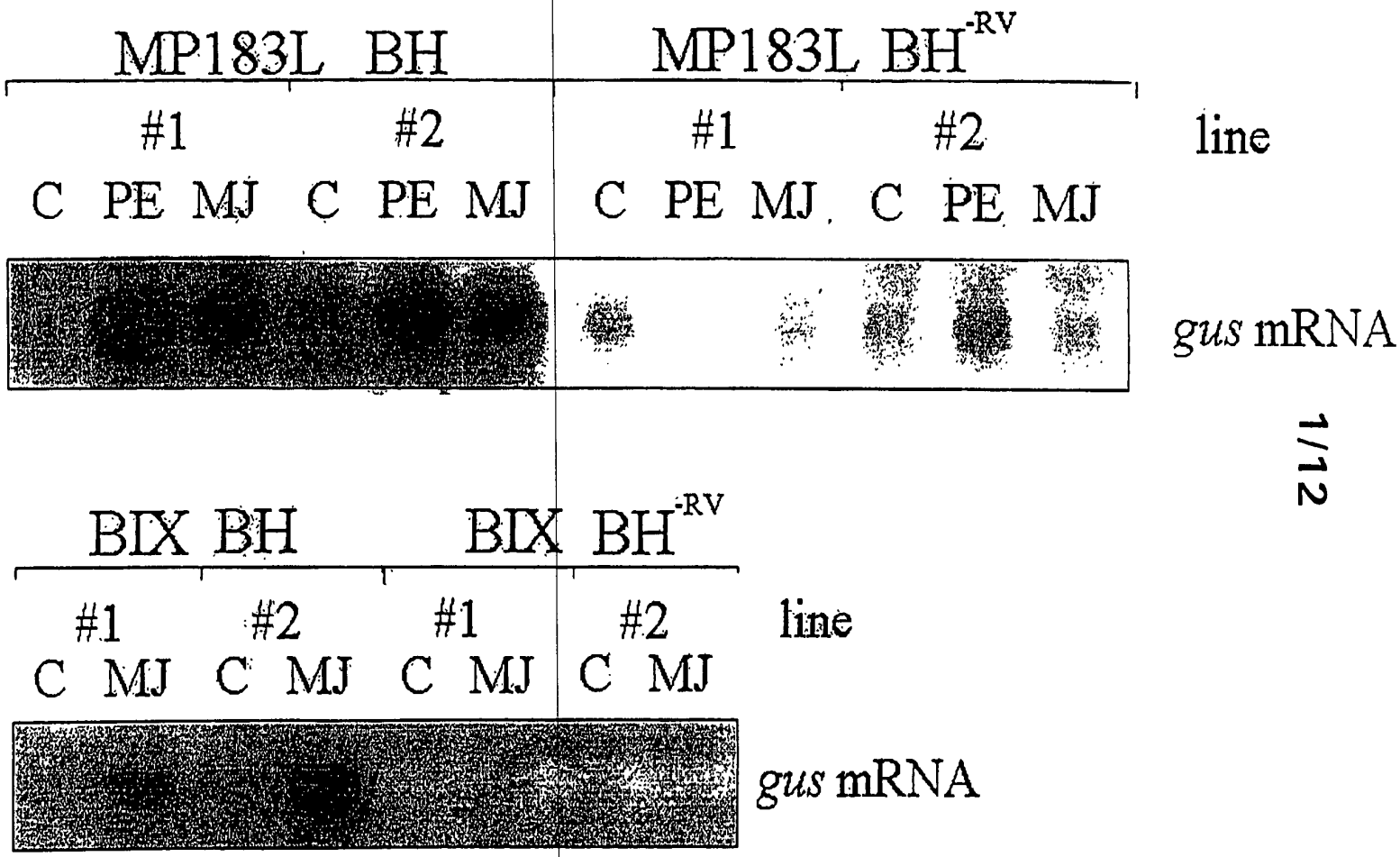


Fig. 1